Studies of user-generated content: a systematic review

Teresa K. Naab, Annika Sehl

Angaben zur Veröffentlichung / Publication details:

Naab, Teresa K., and Annika Sehl. 2017. "Studies of user-generated content: a systematic review." *Journalism: Theory, Practice & Criticism* 18 (10): 1256–73. https://doi.org/10.1177/1464884916673557.

Accepted for publication in *Journalism: Theory, Practice and Criticism*

Studies of user-generated content: A systematic review

Teresa K. Naab*, Annika Sehl

* Teresa K. Naab
Department of Media, Knowledge, and Communication
University of Augsburg
Universitaetsstrasse 10
86159 Augsburg
Germany
Email: teresa.naab@phil.uni-augsburg.de

Abstract

This article presents a review of communication research on user-generated content with a special focus on studies which include a content analysis. The trends of research on this comparatively new and rapidly developing subject are systematically discussed and desiderata are identified. The evaluation is based on a content analysis of pertinent approaches in nine relevant international peer-reviewed journals published from 2004 to 2012. From the results, the article concludes that user-generated content is approached by scholars from a variety of perspectives and offers scope for interdisciplinary cooperation but also notes that several of the challenges posed by the continuously changing nature of the content are not fully met.

Keywords: Content analysis, methods, produsage, systematic review, user-generated content

Studies of user-generated content: A review

Driven by technological developments and rapid user acceptance, user-generated content (UGC) has become a central subject of examination in communication studies. The research on UGC produces findings on a phenomenon that is still relatively new compared to research on the traditional mass media content produced by professionals but has already had a decisive impact on the communication landscape. By adopting different perspectives and using different methods, a number of fields are engaging with the analysis of media participation. These fields include, but are not limited to, journalism studies, which pay attention to the influences of lay producers on traditional media content and the integration of UGC in their products. Other fields are media sociology, media content research, and reception and effects research. An integrative view of the phenomenon of UGC that includes manifold perspectives on the subject makes it possible to observe opportunities for interdisciplinary cooperation and

indicates the borders and intersection of journalism research with other academic areas.

The centerpiece of this article is a systematic review of the empirical communication research on UGC published in international journals. The cumulative process of achieving scientific progress requires researchers to refer to the existing knowledge and improve the methodological deficiencies, check the results, and fill the research gaps. Systematic reviews of research can make a contribution in this regard by providing an overview of the scope of existing research, the prevalence of the procedures used, and the problems identified. The objective is to adopt a systematic approach in order to reveal the focus of the research on a relatively new and rapidly developing research object and in order to determine the priorities set within the scientific system. Systematic overviews and meta-analyses of academic publications have a long tradition in communication studies. They review the scope of the research topics (e.g. Borah, 2011; Bryant and Miron, 2004; Neuman and Guggenheim, 2011), the

distribution of the methods and analytical procedures (Cooper et al., 1993; McMillan, 2000; Trumbo, 2004), and the presentation and reporting standards (Lombard et al., 2002; Lovejoy et al., 2014). They also review the general trends and conceptual deficiencies in scientific publications (Kamhawi and Weaver, 2003). In recent years, several authors have conducted systematic reviews of publications on the Internet and emerging technologies (Borah, 2015; Cho and Khang, 2006; Kim and Weaver, 2002; Peng et al., 2012; Tomasello, 2001; Tomasello et al., 2010) and specific services (e.g. on social networking service (SNS), Zhang and Leung, 2014). This article continues this tradition and presents a systematic review of studies on UGC presented in scientific journals.

The article has two areas of focus: First, it systematically investigates the priority topics and research interests that are examined in relation to UGC and gives an overview of the methods applied in UGC research. Second, the article is especially interested in reviewing studies that apply standardized content analyses to UGC.

The following section will give a broad definition of UGC as applied during this systematic review. The objects of the review will be subsequently outlined. The article will then present the method and findings of the review and draw conclusions about the dynamics of academic research on UGC.

UGC

Since the end of the 20th century, intelligent web services based on new technology have enable users to contribute to media content and interact with other users. As part of this development, platform operators have moved to a central position. They do not produce media content, but provide users with the means to produce and collaborate on content as well as the means to distribute, customize, and develop it. Authors have referred to this phenomenon by a multitude of terms, including social media, Web 2.0, participatory web, UGC, and others (e.g. Bruns and Schmidt, 2011; O'Reilly, 2005; Organisation for Economic Co-operation and Development (OECD), 2007). They have emphasized different aspects,

applications, and operators, and have expressed various hopes for the individual and society (for an overview, see Dahlberg, 2011). Moreover, the fluidity of the subject adds to the challenge of agreeing on a definition. In this systematic review, we will refer to the term ‘user-generated content’ and define it according to the following criteria:

1. *UGC is characterized by a degree of personal contribution*. The users contribute the content themselves; simply receiving or forwarding content and similar activities do not qualify. Such contributions can consist of commenting within the framework of the existing services (e.g. online letters to the editor, user comments to online articles, or comments on blogs), researching and preparing information (e.g. *Wikinews*), and uploading individual text, images, and audio (e.g. blog posts, forum posts, and photo and video platforms). The OECD (2007) refers to this criterion as ‘creative effort’ (p. 8). In the interest of the analysis, only such content will be included that (at least theoretically) could

give a rationale. When turning to applications that can be used to produce and distribute UGC, we will concentrate on those that allow comments and feedback to enable interaction. Applications that simply allow the expression of a judgment without any opportunity for reasoning were excluded because the creative effort of the individual participant is limited. However, it should be kept in mind that the aggregation of such individual social endorsements has new implications which go beyond the effort of the single user, for example, lists of most read articles.

2. *UGC must be published*. In order (at least theoretically) to enable an overall discussion across society as a whole or within a group, it must be accessible to the public or a group (OECD, 2007). In this article, this criterion also applies in principle to content produced on social network sites and weblogs, even though some of these use access restriction with regard to their content. However, bilateral communication is excluded, for

example, content transmitted via applications such as e-mail or instant messaging.

3. *UGC is created 'outside the realm of a profession and professional routines' (OECD, 2007: 8).* Nip (2006) asserted that a distinction can be drawn between analyses of participatory journalism content (e.g. Singer et al., 2011) – and citizen journalism content (e.g. Allen and Thorsen, 2009). According to his definition, participatory journalism are those scenarios where the media seek a degree of proximity to their audience by including the participation of its members in the editorial processes and in the production of publications or broadcasts. In contrast, in citizen journalism, the design-making and decision-making powers are completely in the hands of amateurs.

A stable definition of UGC is a necessary precondition for understanding the changes in the research on the phenomenon over time and between different research areas and theoretical and methodological traditions. This relatively broad definition

allows for the inclusion of research on many different aspects of UGC, some of which have found interest in many different areas of communication science and some of which are more specific, some of which came up during the initial years of amateur participation in online media and some of which have developed later. The systematic review will select studies on UGC by selecting studies that refer to *applications* which allow the production, distribution, and processing of content that fulfills these criteria. Whether the applications allow for participatory journalism or citizen journalism in accordance with the above-mentioned differentiation does not influence their inclusion in the systematic review. Among these applications are spaces for online comments, online letters to the editors, and amateur reports. Such applications are mostly but not exclusively provided by media institutions and complement its professional content. Discussion forums, Usenet, newsgroups, and mailing lists provide the opportunity to post messages and reply without reference to previous professional media content and mostly outside the realm of a

media institution. The same holds true for weblogs, although these tend to cover a narrower range of subjects. Weblogs tend to be published by individuals or small groups of authors, but allow for interaction despite being static, traditional websites. Microblogging systems like *Twitter* feature very short posts. Wikis allow for the collaborative modification of structure and content, even content that was originally produced by others. Social networking applications integrate many of the aforementioned services. They focus on features to construct a user profile and connect it to others' profiles. Photo and video communities gather around uploading pictures and videos. They include many features of the aforementioned services.

Objects of the systematic review

General review of studies

R1. Which applications of UGC are examined over time?

As defined in this article, UGC includes content that is produced, distributed, and processed on a variety of applications. However, not all of them are equally well

regarded by communication researchers. The systematic review aims at outlining how the research has developed over time. It includes research on the technological aspects of the defined UGC applications, on its producers and production context, on its content and design, and on its audience and effects.

R2. In which theoretical context does communication research examine UGC?

UGC services can serve as an object of examination for communication scholars in many fields. A prominent area may be their production context, their relationship with and influence on traditional journalism, their content and their design. However, since the borders between communicator and recipient are continuing to vanish in new media technologies, the audience is of equal importance. The systematic review thus also includes the audiences' perceptions and attitudes towards the applications, usage practices, what they draw from it and the applications' relevance in society. In order to systematize this cursory

overview, to present the quantitative relevance of the different fields, and to identify less prominent research backgrounds, the empirical survey addressed the research contexts in which UGC is examined.

R3. Which designs and methods are used to examine UGC?

On the one hand, methodological approaches, research designs, and measurement clearly depend on the exact object of examination and on the research question. On the other hand, methodological choices also determine the scope of the possible results. The quantitative review will provide an overview of the most frequently applied methods and research designs.

Review of the procedures of the content analyses

R4. Which kinds of content are examined with respect to UGC?

The method of content analysis promises to produce valuable insights when the media output resulting from user participation is examined. Content analyses of the traditional media services primarily show the kinds of content with which

recipients come into contact which may allow conclusions regarding the possible effects to be drawn. Content analyses of UGC additionally give access to the produsers' (on this term, see Bruns and Schmidt, 2011), actual output, their (published) knowledge and opinions, design and composition, that may differ from the professional journalistic forms of presentation. Amateur communicators have an opportunity to draw attention to their own, very individual interests. The result is a wealth of issues extending beyond the portfolio of topics associated with traditional journalism and serving niches outside the taste of a mass audience. The scope for topics among services that can be examined is accordingly diverse.

R5. Which modes of communication does communication research take into account when examining UGC?

In creating online content, amateur communicators can make use of the multimedia nature of the services. They may include text, audio, video, and animated material in their contributions. Furthermore, links and digital references enable the services to offer hypertextuality.

*R*6. Which criteria are used to select UGC for content analyses?

The method of content analysis must continue to develop in tandem with the technological possibilities and adapt to the content to be examined in order to produce valid and reliable findings. Content analysis is one of the central methods employed in communication research, and its critical, careful application and continued development are important tasks within the field: ‘Given that content analysis is fundamental to communication research (and thus theory), it would be logical to expect researchers in communication to be among the most, if not the most, proficient and rigorous in their use of this method’ (Lombard et al., 2002: 587). Although the design of any content analysis is aligned with specific research interests, a number of general difficulties confront scholars conducting online content analysis and the analysis of UGC in particular (McMillan, 2000; Mitra and Cohen, 1999; Schneider and Foot, 2004; Weare and Lin, 2000). One such difficulty is that, because of the highly dynamic nature of the medium, the

online material is not permanent, but highly transient. Consequently, the statistical populations of many online content analyses are not known and are subject to continuous change. UGC compounds this problem because it relies on the constant collaboration of produsers without any stipulated periodicity. In contrast to traditional mass media content, UGC tends to be highly reactive and personalized. Some online content is only displayed at the individual level by a specific user and is dependent on the activities of that user (e.g. his or her friends in a social network). In addition, not all of the content used at the individual level is fully public and easily available for examination. Therefore, researchers must clearly define the sample of their study (for a detailed discussion, see Lacy et al., 2015), explain how they accessed it and how they archived the transient material.

*R*7. To what extent do researchers make use of the opportunities of international comparative studies?

Although it presents challenges regarding sample selection, UGC is digitally available and offers the advantage of

providing easier access to international content, which enables international comparative studies to be carried out.

Method

Sample

The systematic review presented in this article is based on a content analysis of relevant articles in international communication journals. The sample consists of articles published in nine international journals that select manuscripts on the basis of a peer-review process. The sample includes three core journalism studies journals, but also others with a broader scope that publish articles on UGC, including but not limited to journalism studies, in order to observe the borders as well as different perspectives on the topic: *Communications – The European Journal of Communication, European Journal of Communication, Communication Research, Journalism: Theory, Practice & Criticism, Journalism Practice, Journalism Studies, Journal of Communication, Journal of Computer-*

Mediated Communication, and *New Media & Society*.[1] The fact that the study is limited to these journals naturally means that the findings depend on the scope of these journals. The exclusion of other publication forms, such as articles in edited volumes, also represents a limitation in relation to the significance of the systematic review. Nevertheless, scientific journals are acknowledged to be the 'barometer of research trends and reflect the evolution of communication research' (Kamhawi and Weaver, 2003: 7), which makes them particularly relevant for meta-analyses.

The systematic review covers journal issues from 2004 to 2012. Undisputedly, this is a limited time span and thus covers only part of the UGC research. This time span covers a period in which research on UGC started to pick up pace and spread into communication science more broadly. Peng et al. (2012), based on a keyword search in the Social Science Citation Index

[1] Findings on a smaller international journal sample and on a less extensive time period have been published in two book chapters in German language (Naab and Sehl, 2014; Sehl and Naab, 2014). Instead, those publications included German journals that are not part of the analysis presented here.

and the Arts & Humanities Citation Index, state that the term Web 2.0 began to emerge in 2005. However, the development of UGC dates back to the last century and also attracted scientific attention at that time. Tomasello (2001) provides a systematic review of Internet research conducted before the year 2000, which also includes academic attention to mailing lists, newsgroups, and bulletin boards, among others. Additionally, this frame is quite extensive compared to other reviews of publications on Internet research (for an outstanding exception and samples with more limited scope, see Borah, 2015; Tomasello et al., 2010).

Coding process

The analysis of the research published in these journals was conducted in three steps. All of the original articles were assessed to see whether they reported on empirical, primary or secondary analytical studies of UGC in accordance with the definition. The above mentioned list of UGC applications which can be expected to include the defined content led to the

selection of the articles. In addition to manuscripts that examined the aforementioned UGC applications, the analysis included articles that did not examine a specific UGC application but instead examined UGC in general – for example, an article in which professional communicators were asked about their opinions on content produced by amateurs. The analysis also took into account articles that investigated professional media content in regard to the integration of UGC applications, such as examinations of news websites regarding their presentation of comment spaces. Furthermore, the analysis was open to additional applications or types of UGC that were not included in the initial list. Only those articles evaluated as relevant in this first step were included in the systematic review. The decision was made on the basis of the abstract. Six student coders with experience in empirical social research selected the articles. They received comprehensive training. An intercoder reliability test, which included 48 articles, indicated a high level of reliability for the collection

of the articles. The Holsti value was 0.9. In total, 239 articles were selected for further analysis.

The collected articles were analysed more extensively in the second step of the content analysis. The publication date, the publishing journal, and the theoretical research context were coded at the level of the entire article. Every study presented in each article was then individually analysed. Separate methods sections, different applied methods, and new sampling procedures indicated whether an article presented more than one study. Each study was coded with regard to which UGC application was investigated. Multiple coding was possible. In addition, information on the research design and the applied method was collected. The analysis was conducted by seven coders after they had completed comprehensive training. An intercoder reliability test involving 4 percent of the articles indicated a high level of reliability. The Holsti value was between 0.7 and 1.0 for all the variables.

Finally, all of the studies for which a quantitative content analysis was conducted were investigated further. Two coders

recorded the following details: which media content was considered in the published study, which criteria were reported for choosing the respective sample, and whether the design included a comparison between countries or languages. In the third step of the systematic review, the Holsti value was between 0.7 and 1.0 for all of the variables.

Results

In accordance with the selection criteria of the systematic review, all of the 239 manuscripts presented at least one empirical study. The authors of 199 manuscripts limited themselves to the presentation of one study, whereas 36 manuscripts featured two studies, two manuscripts presented three studies, and two further manuscripts presented four studies. Consequently, the content analysis of the variables related to the studies is based on a total of 285 studies.

The *Journal of Computer-Mediated Communication* led the list of publications on UGC, with 74 publications, followed by *New Media & Society* (62), *Journalism: Theory, Practice &*

Criticism (22), *Journal of Communication* (21), *Journalism Practice* (19), *Journalism Studies* (16), *European Journal of Communication* (10), *Communication Research* (9), and *Communication* (6), which put less of a focus on the research topic under consideration.

Examined applications and development over time (R1)

Around a quarter of the empirical observations were related to weblogs (28% of the 285 studies) and discussion forums, Usenet, newsgroups, and mailing lists (22%). Next came social networks (17%). UGC in general, without any closer specification of the actual application, was investigated in 11 percent of the studies. The other UGC applications only received moderate attention.

Although several UGC applications like discussion boards, weblogs, and wikis existed before 2004, these phenomena received very little scientific attention in the first year included in the review (seven published articles in the respective journals in 2004). However, we see a notable increase in the

level of scientific attention they attracted in 2005. Up to 2009, the examined journals published between 19 and 26 relevant articles every year. We see another significant increase in the number of publications in 2011, when 52 articles were published, and almost as many were published in 2012 (49 articles). The early years were dominated by the analysis of discussion forums. Since 2007, weblogs have gained particular significance. Social network sites have become significant for the scientific community since 2009 and were the subject of nearly 40 percent of the studies in 2012. This distribution reflects the development and pervasiveness of the applications. The time lag between the introduction of specific applications and the publication of the research findings is conspicuously large. Although the first weblogs emerged in the late 1990s (Blood, 2000) and were widely disseminated in the early 2000s (Sifry, 2004), it took several years for this development to be reflected in the scientific publications.

Research contexts (R2)

The theoretical context of each article was encoded with a maximum of three context attributes (Table 1). The list of contexts was deduced prior to the analysis and based on similar systematic reviews of academic publications (among others, Cho and Khang, 2006; Kim and Weaver, 2002). Almost half of the manuscripts addressed the research on the processes of produsage. Although it might seem obvious that a systematic review of studies on UGC should find produsage to be the most significant research context, it might also be considered necessary to explain why not all the studies addressed clearly focus on this frame of reference. Produsage was encoded as the theoretical background of an article when what the authors were examining 'is no longer simply usage or production, but something else altogether: produsage' (Bruns and Schmidt, 2011: 4). As Bruns and Schmidt (2011) explained, the terms 'product' and 'usage' may no longer be applicable when the targeted behavior is a collaborative and continuous contribution of content. However, when the researchers were

primarily investigating the usage of online content produced by amateurs, the article was coded in relation to media reception. If the manuscript elaborated on media content theory (e.g. news values), it was coded as media content, and so on. A third of the articles examined UGC services from the perspective of media sociology. They addressed empirical studies concerned with the public, public opinion, integration, group formation processes that take place by means of UGC, or the effects on social norms and values. Almost as frequently, the authors were interested in aspects of interpersonal communication and the characteristics of computerized or direct communication between persons (29%). About a quarter of the articles showed an interest in journalism research (25%) (for a more detailed description of the topics, see below) or political communication (23%). Nineteen percent of the articles were aimed specifically at researching media content (see below, for more extensive information on the analysed content). The reception of participatory services, questions regarding media selection, the processes that occur while

experiences are taking place, such as entertainment by means of UGC (14%), and media effects – that is, the consequences of usage on areas such as impression formation (14%) – were also areas that featured in around a sixth of the articles. Other research areas appeared far less frequently (Table 1).

Table 1

Research contexts

	n	*%*
Produsage	111	46
Media sociology	77	32
Interpersonal communication	68	29
Journalism research	59	25
Political communication	55	23
Media content research	45	19
Media effect	34	14
Media reception	33	14
Diffusion of new media	23	10
Media and communication ethics	15	6
Health communication	12	5
Intercultural communication	13	5
Gender research	11	5
Media theory	5	2
Media economy	5	2
Media and communication law	3	1

Media policy, media system	2	1
Organizational communication, Public relations, advertising	2	1
Visual communication	2	1
Media pedagogics	1	0
Methods, measurement	2	1
Media and communication history	0	0
Nonverbal communication	0	0

N = 239 articles.
Multiple coding was possible.

The emergence of UGC has a distinctive influence on journalism. Its consequences for and differences to traditional media production have often been the subject of debate. Therefore, a more detailed description of articles dealing with journalism research seems relevant. Applications that are specifically examined under the perspective of journalism research are online comments, online letters to the editors and amateur reporters. SNS are researched but to a lesser extent than in other research contexts (e.g. media reception). Most of the studies (62%) addressed UGC outside institutional journalism. In contrast, less than a quarter (23%) closely

examined audience participation within institutional journalism and thereby a professional journalistic framework. Three percent of the studies examined both participatory journalism services and citizen journalism services. In 12 percent of the studies, there was no information on the production context. A closer look at the articles reveals two broad topics, each of which amounts to about half of the sub-sample: The first category contains articles on the relationship between UGC outside the institutional media and professional journalism, a typical article may refer to ‘Competition, complementarity or integration. The relationship between professional and participatory media’ (Neuberger and Nuernbergk, 2010). Several of the articles in this category deal with weblogs (e.g. Gil de Zúñiga et al., 2011; Messner and Distaso, 2008; Reese et al., 2007; Vos et al., 2012). Also, SNS like *Facebook* and *Twitter* are discussed in this respect, albeit less frequently (e.g. Verweij, 2012). Other subtopics occur on a more singular basis. The second category contains articles on how professional journalism is integrating the audience. A

typical article may refer to ‘Participatory journalism practices in the media and beyond: An international comparative study of initiatives in online newspapers’ (Domingo et al., 2008), which describes audience participation in different news outlets, different sections like science journalism or different countries (see, for example, also Artz and Wormer, 2011; Deuze et al., 2007; Secko et al., 2011).

Choices of design and method (R3)

Half of the studies used an open research approach, whereas a quarter posed explicit research questions (25%) or formulated hypotheses (25%).[2] Similar results can be found in other systematic reviews on methodological approaches in communication research (Cho and Khang, 2006; Cooper et al., 1993; Trumbo, 2004). Interestingly, there was no significant

[2] A study was coded as providing a research question, when a research question was explicitly stated in the text. A sole indication of interest in specific results embedded in the theory section or the like was not coded as a research question. Similarly, only explicitly stated hypotheses were coded and not general assumptions made in the text.

increase in the testing of hypotheses over the course of time. There was no chronological trend towards posing explicit research questions and hypotheses for any of the applications (see also Cho and Khang, 2006).

The majority of the reviewed studies applied quantitative-standardized methods to examine UGC (Table 2).

Table 2

Applied methods

	n	*%*
Standardized content analysis	79	28
Standardized survey	79	28
Non-standardized content analysis, text analysis, discussion analysis	68	24
Non-standardized interviews	32	11
Standardized observation	11	4
Non-standardized observation	5	2
Other methods	11	4
Total	285	100*

N = 285 studies.
*Deviations from 100 percent are due to rounding.

Standardized content analysis and standardized surveying were the most frequently applied methods, followed by qualitative text analyses and non-standardized interviews. The dominance of quantitative methods in communication studies

in the examination of UGC bears out (albeit less explicitly) the results of previous systematic reviews of methods used in the examination of Internet studies (Peng et al., 2012; Zhang and Leung, 2014) and other topics (Cooper et al., 1993; Kamhawi and Weaver, 2003; Trumbo, 2004). However, when reviewing research on emerging technologies, Borah (2015; also Cho and Khang, 2006) finds quantitative and qualitative approaches to be spit about equally.

Articles could include more than one study: around every 10th article presented a combination of methods (12%). The results of earlier reviews in other subject areas varied considerably in relation to the combination of methods (Kamhawi and Weaver, 2003; Trumbo, 2004), and surely depends on the selection of the journals.

The majority (75%) of the 285 studies examined UGC in cross-sectional analyses. Nevertheless, 11 percent were case studies. Eight percent of the authors reported on experiments (for similar results, see Cooper et al., 1993; Kamhawi and Weaver, 2003; Trumbo, 2004; Zhang and Leung, 2014). Longitudinal

observations and panels were found in 5 percent of the studies.[3]

Detailed review of the content-analytic studies of UGC (R4, R5, R6, R7)

The present systematic review places a special focus on standardized content analyses. In total, 79 of the published studies applied this method. Of these, 12 examined professional websites in regard to their inclusion of UGC formats. These were not examined further because these studies themselves merely observe which participatory features were offered by professional journalism to enable users to participate. They do not allow for any findings on the

[3] Cross-sectional studies involve the analysis of data collected at one specific point in time, while longitudinal studies collect data on more than one point in time and analyse trends over time. Panel studies are a particular form of longitudinal study in which the same measures are repeatedly collected from the same sample at different points in time. A study has been coded as presenting a case study when it described one or few research objects by empirically conducting data on this object or analysing secondary data. However, it did not intend to draw conclusion on other objects in the same category of user-generated content (UGC) and the research object is not supposed to stand as an example for similar objects.

methods and measurement decisions in UGC analysis as traditional websites, rather than UGC platforms, were examined.

Communication studies reveal a clear preference for political UGC (Table 3).[4]

Table 3
Topics examined

	n	*%*
Politics	25	37
Crime	4	6
Health	3	5
Accidents	3	5
Research, education, science	1	2
Social issues, religion	1	2
Private issues	1	2
Economy	1	2
Multi-topic related	19	28
Other topics	7	10
Not recognizable	2	3
Total	67	100*

N = 67 studies conducting quantitative content analyses.
*Deviations from 100 percent are due to rounding.

[4] A list of possible content categories was deduced prior to the analysis based on Quandt (2008). An "other" category included further topics that were not previously included in the list.

Thirty-seven percent of the 67 content analyses examined political issues. Other UGC was analysed far less frequently. Although social media clearly provide amateur communicators with the opportunity to make their interests public and to communicate their private opinions, communication research evidently prefers to focus on messages with a societal orientation. Twenty-eight percent of the content analyses addressed services that allow open communication on different topics. The content of what was being examined presumably is less important for these research approaches than other characteristics of amateur communication or, alternatively, they explicitly sought a comparison between UGC with different focuses.

The majority of the content analyses concentrated on the analysis of texts (Table 4).

Table 4

Kinds of content examined

	n	*%*
Text	60	90
Links	18	27
Moving image	7	10
Fixed image	4	6
Audio, music	2	3
Advertising	1	2
Animation	0	0

N = 67 studies conducting quantitative content analyses.
Multiple coding was possible.

Multimedia elements, such as moving and fixed images, audio material, and animations provided by amateurs, were rarely examined. However, the hypertextuality of the online services was taken into account through the analysis of links.

The electronic availability of UGC makes it easier to access international material, which is troublesome and costly to acquire when analysing traditional media. Content analysis studies that compared countries were nevertheless in the minority (9%).

Due to the above-mentioned challenges related to defining the statistical population and sampling in UGC analyses, the

systematic review registered which criteria guided the selection of the examined content (Lacy et al., 2015). Most authors limited their analyses to specific time periods (80% of the 67 content analyses), UGC services with a certain thematic focus (78%), or specific language territories or countries (75%) – often without providing reasons for so doing.[5] Besides focusing the research interest on a specific time, issue, and cultural aspects, researchers must decide how to sample the actual cases. The authors of 51 percent of the content analyses stated that they selected services because of their wide reach (e.g. services with a high number of visits or page impressions, high advertising revenue, etc.). This indicated a conscious selection process, oriented towards the potential circle of those who perceive and are potentially influenced by the content. Niche services for small, fragmented target groups, however, were consequently excluded. Alternatively, authors selected UGC applications because of their inclusion in meta-data

[5] Since research can be limited to specific time, thematic, and territory frames at the same time, multiple coding was possible.

indexes, such as editorial lists, search engine-generated lists, and collaborative indexes (e.g. social bookmarking lists) (39% of the content analyses). Whether results based on such selection can be generalized naturally depends on the quality of the indexes used, most of which are incomplete. Furthermore, researchers can sample amateur communicators and analyse the UGC produced or used by this group. This selection strategy was used in only 12 percent of the content analyses.[6]

Discussion

The present systematic review provides an initial impression of the dynamics of communication research on UGC. However, any interpretation must take into account the restrictions of the study. First, the findings of the review can only present the distributions as they appear in the journals that were included. However, any process of selecting journals has

[6] Since research can draw on the criteria of wide reach, inclusion in meta-data indices, and selection of users at the same time, multiple coding was possible.

unavoidable limits. The inclusion of additional publications in the future is most likely to gradually change the image that has emerged here. Second, UGC is a field of continuous and fast technical, economic, and social change. This review covers only a limited time period of research on UGC. It does neither picture this most early start nor the most recent status quo. However, it may be helpful in summarizing the period in which research on UGC grow to an important factor in communication science and helpful in analysing change in research topics in general. Third, the framework of the examination is significantly determined by the review's definition of UGC. This is why the study chose quite a broad definition that includes most applications generally assumed under the term UGC.

Most publications analysed UGC with respect to produsage – that is, the design process from the perspective of amateurs in general. They focus on the continuous co-creation and usage of non-institutional communicators. In other words, within the theoretical frameworks of their empirical investigations, the

authors adapted their approach to address the new phenomenon of the intersection between production and usage that is unique to UGC. The broad relevance of UGC for many different scholarly fields is demonstrated by the fact that researchers also place their analysis of UGC in the context of political communication research, media sociology, and reception studies as well as traditional journalism, interpersonal communication, and many others. In this vein, Peng et al. (2012) show that Internet research in general is not divided in line with disciplines. What is special about UGC is precisely the link between the questions associated with research on communicators and those associated with research on recipients. This not only raises new research questions but also opens up the possibility and the need for interdisciplinary cooperation.

The empirical analyses mainly addressed weblogs, discussion forums, social network sites, and online comments. All of the remaining applications were only marginally examined. This is probably due to the fact that weblogs and forums are older

applications, having had more time to generate research interest. It can be assumed that the number of publications with respect to social networks will continue to increase. In comparison to social networks, however, weblogs and forums have a relatively linear structure that allows greater access for analysis. Most studies engaged with UGC outside a professional editorial framework. Less than a third addressed the question of participation in what are considered, in this context, to be traditional media. This finding also corresponds to the reality of the situation in which the majority of participation takes place independently of professional editorial media services. Nevertheless, these citizen journalism services can also be a driving force for participation in professional editorial journalism. Even so, in 12 percent of the studies, it was impossible to identify whether the authors were investigating participatory journalism applications or citizen journalism applications. This makes the interpretation of the findings and the comparison with other studies difficult, and

must be regarded as a deficit in the presentation of the research objects.

Regarding the methodological approach, the research showed a preference for quantitative-standardized methods in the studies on UGC in the selected journals. Thus, while hypothesis-based tests of theory are limited, scholars do not engage in open exploration but focus on generalizable, quantitative results. This orientation is not specific to social media research, but can also be found in comparable reviews (Cooper et al., 1993; Kamhawi and Weaver, 2003; Trumbo, 2004). Content analyses and surveys were the most frequently applied methods, followed by qualitatively orientated text analyses and qualitative interviews. The content analyses evidently drew on the potential directly to examine the knowledge and opinions communicated by media users in order to reduce the distorting effects that can appear in surveys. Interestingly, the number of methods that involved the testing of hypotheses did not increase over time, as might have been expected. Evidently, many studies adopted an exploratory

approach to dealing with new applications. However, such an exploration with open research designs is not followed by more theory-driven and empirically testing designs.

Technological development has changed communication on the Internet. Professionals as well as lay communicators can draw on textual, audio, video, and animated styles to express themselves and interact with others, and UGC applications make it possible to refer to, link with and integrate the contributions of others. These new modes of communication constitute, to some extent, the particularities of UGC and influence the usage and content of the applications. However, the review shows that the potential inherent in online media and UGC is not yet being fully exploited. The authors tended to fall back on classic text analyses and neglect multimedia modes of communication. Links were the only notable exception accounted for in the investigations. Auditory and visual elements, meanwhile, were rarely included. As such, the research was unable to reflect the potential methods of expression offered by online media. It is doubtful whether the

studies were able to account properly for the produsers' expressions because these use the multimedia sign systems as a normal aspect of their communication.

Aside from providing the users with additional modes of communication which can be researched, the technological development confronts the social science methodology with new challenges regarding the analysis of online content and of UGC in particular. UGC is highly dynamic in nature. Lay producers contribute intermittently. Furthermore, the content is reactive, personalized, and partly private. This creates challenges regarding the selection of UGC that should be recognized in the sampling process. The authors of the content analyses only rarely take amateur communicators into account as a selection basis (i.e. by selecting people and asking them to make available the social web content that they had used and created). The limited use of this procedure means that the specific surfing practice of the users rarely guided the selection of the analysed content. The studies could, therefore, only account for the reactivity and personalization of UGC to a

limited extent. Consequently, it is doubtful whether they were able to examine the material in its full scope and with an eye to the particularities that distinguish it from traditional media content. Furthermore, the potential use of UGC content analysis as a complementary method to reactive survey findings is not being explored.

Although online media simplify access to transnational material, the amount of international comparative studies is limited. The review cannot clarify if this is due to practical research considerations and the significant extra effort associated with an extensive range of material and cultural and language barriers in its examination. Also, the research questions probably do not aim at the analysis of transnational material. Comparative studies may constitute an interesting field for future research. The dominance of cross-sectional studies and the lower number of longitudinal and experimental studies is unsurprising with such a relatively new research object. However, these observations highlight potential areas for future study.

Conclusion

Overall, the systematic review has shown that UGC is a research object that is compatible with many areas of communication research and is adopted by scholars in a variety of fields. Nonetheless, UGC offers scope for a more detailed examination. One can presume that the examination practices, which seem to be oriented towards the content analysis of traditional media in numerous ways, are the consequence of an incomplete debate on methodological standards and the lack of established transferable examples. This systematic review seeks to contribute in small part towards overcoming these challenges by providing an overview of the current practices. Its results clearly depend on the sample of journals included, and an ongoing review of this rapidly developing subject will prove fruitful.

References

Allen S and Thorsen E (2009) *Citizen Journalism: Global Perspectives*. New York: Peter Lang.

Artz K and Wormer H (2011) What recipients ask for: An analysis of ‘user question generated’ science coverage. *Journalism* 12: 871–888.

Blood R (2000) Weblogs: A history and perspective. Available at:

http://www.rebeccablood.net/essays/weblog_history.html

Borah P (2011) Conceptual issues in framing theory: A systematic examination of a decade’s literature. *Journal of Communication* 61: 246–263.

Borah P (2015) Emerging communication technology research: Theoretical and methodological variables in the last 16 years and future directions. *New Media & Society*. Epub ahead of print 17 December. DOI: 10.1177/1461444815621512.

Bruns A and Schmidt JH (2011) Produsage: A closer look at continuing developments. *New Review of Hypermedia and Multimedia* 17: 3–7.

Bryant J and Miron D (2004) Theory and research in mass communication. *Journal of Communication* 54: 662–704.

Cho CH and Khang HK (2006) The state of Internet-related research in communications, marketing, and advertising: 1994–2003. *Journal of Advertising* 35: 143–163.

Cooper R, Potter WJ and Dupagne M (1993) A status report on methods used in mass communication research. *Journalism Educator* 48: 54–61.

Dahlberg L (2011) Re-constructing digital democracy: An outline of four 'positions'. *New Media & Society* 13: 855–872.

Deuze M, Bruns A and Neuberger C (2007) Preparing for an age of participatory news. *Journalism Practice* 1: 322–338.

Domingo D, Quandt T, Heinonen A, et al. (2008) Participatory journalism practices in the media and beyond: An international comparative study of initiatives in online newspapers. *Journalism Practice* 2: 326–342.

Gil de Zúñiga H, Lewis SC, Willard A, et al. (2011) Blogging as a journalistic practice: A model linking perception, motivation, and behavior. *Journalism* 12: 586–606.

Kamhawi R and Weaver D (2003) Mass communication research trends from 1980 to 1999. *Journalism & Mass Communication Quarterly* 80: 7–27.

Kim ST and Weaver D (2002) Communication research about the Internet: A thematic meta-analysis. *New Media & Society* 4: 518–538.

Lacy S, Watson BR, Riffe D, et al. (2015) Issues and best practices in content analysis. *Journalism & Mass Communication Quarterly* 92: 791–811.

Lombard M, Snyder-Duch J and Bracken CC (2002) Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research* 28: 587–604.

Lovejoy J, Watson BR, Lacy S, et al. (2014) Assessing the reporting of reliability in published content analyses: 1985–2010. *Communication Methods and Measures* 8: 207–221.

McMillan S (2000) The microscope and the moving target: The challenges of applying content analysis to the World Wide Web. *Journalism & Mass Communication Quarterly* 77: 80–98.

Messner M and Distaso MW (2008) The source cycle: How traditional media and weblogs use each other as sources. *Journalism Studies* 9: 447–463.

Mitra A and Cohen E (1999) Analyzing the web: Directions and challenges. In: Jones S (ed.) *Doing Internet Research: Critical Issues and Methods for Examining the Net*. Thousand Oaks, CA: SAGE, pp. 179–202.

Naab TK and Sehl A (2014) Inhaltsanalytische Untersuchungen von User-Generated-Content-Angeboten: Eine Bestandsaufnahme zur Anwendung der Methode [Reviewing content analyses of user-generated content applications]. In: Sommer K, Wettstein M, Wirth W, et al. (eds) *Automatisierung in der Inhaltsanalyse*. Köln: Herbert von Halem, pp. 127–144.

Neuberger C and Nuernbergk C (2010) Competition, complementarity or integration? The relationship between professional and participatory media. *Journalism Practice* 4: 319–332.

Neuman WR and Guggenheim L (2011) The evolution of media effects theory: A six-stage model of cumulative research. *Communication Theory* 21: 169–199.

Nip JYM (2006) Exploring the second phase of public journalism. *Journalism Studies* 7: 212–236.

O'Reilly T (2005) What is Web 2.0. Available at: http://oreilly.com/web2/archive/what-is-web-20.html

Organisation for Economic Co-operation and Development (OECD) (2007) Participative web: User-created content. Available at: http://www.oecd.org/sti/38393115.pdf

Peng TQ, Zhang L, Zhong ZJ, et al. (2012) Mapping the landscape of Internet studies: Text mining of social science journal articles 2000–2009. *New Media & Society* 15: 644–664.

Quandt T (2008) Neues Medium, alter Journalismus? [New medium, old journalism?]. In: Quandt T and Schweiger W (eds) *Journalismus online – Partizipation oder Profession?* Wiesbaden: VS Verlag für Sozialwissenschaften, pp. 131–155.

Reese SD, Rutigliano L, Hyun K, et al. (2007) Mapping the blogosphere: Professional and citizen-based media in the global news arena. *Journalism* 8: 235–261.

Schneider SM and Foot KA (2004) The web as an object of study. *New Media & Society* 6: 114–122.

Secko DM, Tlalka S, Dunlop M, et al. (2011) The unfinished science story: Journalist – Audience interactions from the globe and mail's online health and science sections. *Journalism* 12: 814–831.

Sehl A and Naab TK (2014) User generated content im Auge der Kommunikationswissenschaft: Deskription eines Forschungsfeldes [User generated content in the eye of communication studies: Description of a field of research]. In: Stark B, Quiering O and Jackob N (eds) *Von der Gutenberg-Galaxis zur Google-Galaxis. Alte und neue Grenzvermessungen nach 50 Jahren DGPuK*. Konstanz: UVK Verlagsgesellschaft, pp. 117–133.

Sifry D (2004) State of the blogosphere, October. Available at: http://www.sifry.com/alerts/2004/10/oct-2004-state-of-the-blogosphere-corporate-bloggers

Singer JB, Hermida A, Domingo D, et al. (2011) *Participatory Journalism: Guarding Open Gates at Online Newspapers*. Malden, MA: Wiley-Blackwell.

Tomasello TK (2001) The status of Internet-based research in five leading communication journals, 1994–1999. *Journalism & Mass Communication Quarterly* 78: 659–674.

Tomasello TK, Lee Y and Baer AP (2010) 'New media' research publication trends and outlets in communication, 1990–2006. *New Media & Society* 12: 531–548.

Trumbo CW (2004) Research methods in mass communication research: A census of eight journals 1990–2000. *Journalism & Mass Communication Quarterly* 81: 417–436.

Verweij P (2012) Twitter links between politicians and journalists. *Journalism Practice* 6: 680–691.

Vos TP, Craft S and Ashley S (2012) New media, old criticism: Bloggers' press criticism and the journalistic field. *Journalism* 13: 850–868.

Weare C and Lin WY (2000) Content analysis of the World Wide Web: Opportunities and challenges. *Social Science Computer Review* 18: 272–292.

Zhang Y and Leung L (2014) A review of social networking service (SNS) research in communication journals from 2006 to 2011. *New Media & Society* 17: 1007–1024.

About the authors

Teresa K. Naab is a senior researcher at the University of Augsburg, Germany. Her research interests include audience research, computer-mediated communication, and social science methods.

Annika Sehl is a Research Fellow at the Reuters Institute for the Study of Journalism at the University of Oxford, UK. Her research interests are mainly journalism and comparative research.