

SOCIAL MEDIA IN JOURNALISTIC INQUIRY - REQUIREMENTS FOR A SEARCH SYSTEM INCORPORATING JOURNALISTIC QUALITY CRITERIA

Complete Research

Kilian, Thomas, University of Koblenz-Landau, Koblenz, Germany, kilian@uni-koblenz.de

Scherp, Ansgar, Leibniz Information Center for Economics and Kiel University, Kiel,
Germany, a.scherp@zbw.eu

Hammes, Eva, University of Koblenz-Landau, Koblenz, Germany, ehammes@uni-koblenz.de

Abstract

Social media allow journalists to incorporate additional information into their journalistic inquiry. However, journalists often cannot be certain of the quality of information from social media. IS research has yet to sufficiently address this issue. Most notably, there is no information system that supports searching for social media content about a specific topic while considering journalistic quality criteria. This paper presents the results of a design science project, focusing on an extensive literature analysis and on an interview study that was conducted to outline the requirements of this information system. The empirical analysis indicates that, in particular, the actuality, independence and credibility of the sources are relevant quality criteria for the use of social media for journalistic inquiry. The paper concludes with a presentation of the study's limitations and an outlook of further steps in the development of the system.

Keywords: Design science, social media, journalistic inquiry.

1 Introduction

Social media give journalists the opportunity to access new bodies of knowledge that facilitate journalistic inquiry and, ultimately, the creation of better journalistic content. In this paper, social media are understood as Internet-based applications that aim for an integration of users and the creation of user-generated content (Kaplan and Haenlein, 2010). Social media already have an enormous influence on journalism, as evidenced when journalists used YouTube videos and news on Facebook and Twitter to report from the war zone during the 'Arab Spring' (Cottle, 2011).

This instance and other examples illustrate how media companies and individual journalists use social media as an additional channel to publish further content and to better target their recipients. More importantly, these examples show that social media also provide an opportunity to increase journalistic quality based on information that was previously difficult to access.

The main function of journalism is gathering and inquiring, selecting and processing, and presenting and disseminating news and information that are relevant and based on facts (Harcup, 2009). The inquiry of information is one of the most important and time-consuming tasks in journalism. In a study conducted in 2005, German journalists stated that they spent on average 117 minutes per workday on inquiry (Weischenberg et al., 2006). While older studies from Germany found that social media did

not play a major role in journalistic inquiry (Machill et al., 2008), the use of social media for this task is currently increasing. A study conducted in Russia, Poland, and Sweden on the journalistic use of social media found that 67-76% of journalists use social media for the generation of ideas and the search for information, which are two important facets of journalistic inquiry (Appelberg et al., 2014). The study by Gulyas (2013) of British, German, Swedish, and Finnish journalists' social media practices found a high use of social media in all countries. However, the study also characterized social media as an additional instrument for journalistic inquiry (e.g., for sourcing and verifying stories) that complements traditional sources. A recent study showed that in two quality Belgian newspapers, references to social media have become common. Both newspapers published on average five articles per day between 2006 and 2013 referring to Facebook, Twitter, and/or YouTube (Paulussen and Harder, 2014). Altogether, social media support classic inquiry functions from the identification of a topic to fact checking. However, the use of social media during inquiry has led to stronger self-referentiality in journalism (Messner and DiStaso, 2008).

Despite these studies, research has not yet spent much effort on examining how exactly social media are integrated into journalists' daily routines (Paulussen and Harder, 2014). With the quantitative increase of social media content, a qualitative degradation is feared because in social media, traditional editorial quality management is obviously absent. Thus, the requirements regarding the selection of information sources and the critical reflection of the content are raised for journalists. Professional journalists therefore face the challenge of integrating and merging the many common threads that are created on social media into their inquiry. To accomplish this task, a future journalist needs skills in social media, particularly tools from information technology that help to navigate the data flood in social media (Hermida, 2010). Current search engines do not suffice; although initial contributions that address this research gap have recently been published, existing search engines do not explicitly consider the quality requirements that are commonly applied in the context of journalistic inquiry (Diakopoulos et al., 2012; Garbett et al., 2014; Heravi et al., 2014; Schifferes et al., 2014; Zubiaga et al., 2013). In addition, there is research on the credibility of social media, e.g., Twitter. However, this research does not target the precise needs of journalists (Castillo et al., 2011; Qazvinian et al., 2011). In summary, from an IS research perspective, there are currently no robust tools for sourcing and verifying news on social media (Schifferes et al., 2014), let alone a tool that supports active journalistic inquiry in social media. In this context, Sikdar et al. (2013) correctly noted that specific constructs have to be defined to reasonably operationalize credibility and similar quality criteria. This notion applies all the more for the quality of social media for journalistic use. From these considerations, the following research questions are derived:

1. How can the vast magnitude of content in social media be differentiated in terms of its quality and relevance for journalists?
2. Which requirements are needed in an information system to enable inquiry in social media while simultaneously considering journalistic quality?

To date, the first two questions have not been sufficiently considered and answered within the literature. Although there is comprehensive research on the quality of media in communication and media studies, there is a lack on research that differentiates and classifies social media from a journalistic perspective in terms of their quality. In addition, the second research question has not been sufficiently answered. Although there are many articles that address knowledge extraction from social media, there is no approach addressing journalistic inquiry that observes all or at least most forms of social media. Ultimately, the semantic integration of social media sources and a classification of their relevance in media studies serve as a basis of innovative information systems for the future 'data journalism' that examines, evaluates, and uses social media during its editorial activities.

To answer the research questions, a design science approach is applied. In this study, the stage model by Kuechler and Vaishnavi (2008) is adopted, which distinguishes between the stages of Awareness of the Problem, Suggestion, Development, Evaluation, and Conclusion. We focused on the first two stages and structured this study as follows. To define the problem, the current state of the art of journalistic

inquiry and the evaluation of media with regard to their quality are presented. On this basis, expert interviews are used to examine the meaning of social media for journalists, which social media are used, and how their quality is perceived (Awareness of the Problem). By way of example, a usage scenario is presented. The scenario facilitates operationalizing the results of the interviews with regard to the actual requirements of an information system. Based on the derived problem definition, the initial requirements of a system that structures and integrates knowledge from different, heterogeneous social media sources is derived (Suggestion). The goal of this system is to be unsusceptible to the fast-moving and frequently changing current social media content and social media sources. Finally, in line with the summary, the study's limitations and the subsequent steps of development are presented.

2 Definition of the Problem

2.1 Quality of journalism and information quality

In his reflections on the quality of journalism, Bogart (2004) names criteria that are typically mentioned when journalists are asked for quality. Among these criteria are integrity, fairness, accuracy, comprehensiveness, diligence in discovery, authority, breadth in coverage, and variety of content.

This discussion also resonates in academia, where the analysis of the quality of journalism has a long-standing tradition. Merrill (1968) included independence and financial stability, integrity, social concern, good writing/editing, strong opinion and interpretive analysis, staff professionalism and intelligence, the power to influence opinion leaders, the expansion of readers' education and intellect, world consciousness and an emphasis on politics, international relations, economics, social welfare, culture, education, and science as quality indicators. Burgoon, Burgoon and Atkin's (1982) most important criteria were accuracy, depth, impartiality, investigative enterprise, literary style and the sophistication of treatment. Gladney (1990) differentiated organizational and content standards: the latter are strong local news coverage, accuracy, good writing, visual appeal, a strong editorial page, a community focus, news interpretation, a lack of sensationalism and comprehensive coverage. The discussion on news quality also continued for news published online. The most important web news quality indicators by Gladney et al. (2007) were credibility, utility, immediacy, relevance, ease of use, fact-opinion separation, clear paths, simplicity, exclusivity, hyperlocalness, good writing and paramount content. Many other authors provided criteria catalogues to measure journalistic quality (Picard, 2000), stressing in particular their relevance, breadth, credibility, or professionalism. Relevance in the context of journalism concerns the meaning of topics to a variety of recipients. The criterion of breadth means a vast magnitude and variety of content, e.g., with regard to different topics. Professionalism is especially important to assess the neutrality of media coverage and therefore applies to journalists.

In the context of social media, the credibility of political blogs in particular is regularly questioned due to their subjectivity (Sweetser et al., 2008). However, altogether, few studies have referred to the journalistic quality of social media, and if so, they have mentioned it only in passing. For instance, the use of blogs by journalists was found not to influence their work processes and attitudes towards journalistic quality standards. Moreover, it was demonstrated that journalists using blogs (actively or passively) considered the meaning of inquiry and information selection in the future even more important than did journalists who did not use blogs (Welker, 2007). Zeller and Wolling (2010) presented a framework for the content analysis of online content with regard to its quality but did not explicitly consider social media. In their conceptualization of quality, they named aggregated user data such as visits, clicks, and user numbers as relevant indicators for online content analysis. However, they did not empirically test their framework. Neuberger (2012) asked users to evaluate journalistic offerings on the Internet – among them social media such as Twitter, social networks, blogs and video-sharing websites – with regard to their quality. However, the study did not comprise an evaluation according to professional criteria from the point of view of journalists who used social media for their inquiry. Altogether, there is not much robust research on the criteria of the quality of social media, let alone on

quality with regard to journalistic inquiry. Thus, it is necessary to conduct additional empirical studies in this context.

However, it must be noted that the discussion on the quality of journalism is not undisputed. For example, Picard (2000) evaluated the discussion on quality in journalism as problematic because it is not clear what constitutes quality and what elements make up the concept. He criticized criteria lists because most criteria are not measurable or are only indirectly measurable.

This discussion is mirrored by a rich body of studies on information and data quality in IS research, resulting in several different frameworks (e.g., Knight and Burn, 2005). In an IS context, information of high quality is understood as information that fits the needs of users in a particular context, i.e., information that meets information requirements or exceeds them and thereby provides high value to users (Eppler and Wittig, 2000). The often-cited data quality framework by Wang and Strong (1996) is a generic model that has been applied in different contexts, and in some aspects, it is surprisingly similar to models of journalistic quality (see Table 1). Overall, the model consists of 15 quality dimensions that are organized in four categories. The intrinsic quality category captures inherent information quality properties such as believability, accuracy, and objectivity. Interestingly, believability (or credibility) is a criterion that was also mentioned by Picard (2000). The contextual data quality category requires quality in terms of the actual task that is performed with the data. This category captures properties such as relevance, timeliness, and completeness. The representational data quality category is focused on the representation of data, e.g., for interpretability or ease of understanding in order to be effectively processed by data users. Finally, the accessibility data quality category is concerned with both the accessibility of the data and access security. The following table matches the quality categories to the journalistic quality mentioned above in an exemplary way without any claim of completeness.

| Categories by Wang and Strong (1996) | Quality Criteria for Journalism |
|--|--|
| Intrinsic: | |
| - Believability | Also mentioned as credibility by Picard (2000) and Gladney et al. (2007) |
| - Accuracy | Also mentioned by Burgoon, Burgoon and Atkin (1982) and Gladney et al. (2007) |
| - Objectivity | Also mentioned by Burgoon, Burgoon and Atkin (1982) as impartiality and by Gladney et al. (2007) as lack of sensationalism |
| Contextual: | |
| - Relevance | Also mentioned by Gladney et al. (2007) as exclusivity |
| - Timeliness | Also mentioned by Gladney et al. (2007) as immediacy |
| - Completeness | Also mentioned by Burgoon, Burgoon and Atkin (1982) as depth |
| Representational: | |
| - Interpretability and ease of understanding | Also mentioned by Gladney (1990) as visual appeal and by Gladney et al. (2007) as good writing and ease of use |
| Accessibility: | Not represented in research on quality in journalism |

Table 1. Comparison: Data and information quality and quality in journalism

Thus, overall there is a high level of alignment for information quality from IS research on the one hand and criteria for journalistic quality on the other hand. In the following section, we use the framework by Wang and Strong (1996) to present the results of our empirical study (see Table 3).

2.2 Empirical study

Data collection and analysis. For our empirical study, we conducted interviews with experts who use social media. The aim was to gain a deeper understanding of the current usage of social media and to be able to design a system that considers relevant user behavior and, in particular, the needs of users and journalistic quality criteria in the search process.

In the interviews, journalists were asked a) how they assess the relevance of social media for journalistic inquiry and how they apply social media, b) how they assess the quality of this use and c) which typical operational scenarios they anticipate for social media as an inquiry instrument. Altogether, 14 active journalists were chosen as experts possessing comprehensive knowledge and experience in social media. The participants' ages ranged from 25 to 56 years, with professional experience from three to 30 years. Ten of the interviewed journalists were male, and four were female. Eight of them worked in the field of printed media for renowned newspapers such as Sueddeutsche or ZEIT, and two of them worked for local newspapers. One of the journalists worked exclusively on the online editorial team of a television station. Three of the participants worked for print as well as for broadcasting. Many of them focused on politics in a broader sense. Only one of the journalists focused on tabloid stories. The following table, Table 2, further describes the sample. We left out some information mentioned above (such as sex and age) from the table to ensure the anonymity of the journalists.

The interviews were conducted in the summer of 2013 and were recorded via telephone, and they took 19 to 39 minutes. During the interviews, the experts were allowed to answer as they liked. The interviews were transcribed before the content was analyzed using the software program MaxQDA. As a framework for analysis, the aforementioned model by Wang and Strong (1996) was applied. Table 3 gives an overview of social media, their journalistic use and quality evaluations according to different criteria by journalists. The statements below are taken directly from the interview transcripts.

Relevance and usage of social media for inquiry. The participants regarded Twitter and Facebook as particularly relevant social media. Twitter was used by all of the experts and was perceived as beneficial due to the openness of the platform and to the simple structuring and sorting of the content. Moreover, Twitter provides a rapid overview of currently discussed topics with its 'trending topics' feature. Facebook, however, was criticized because, compared to Twitter, it does not provide a simple overview of relevant topics. In addition, following certain people on Facebook is not as easy as it is on Twitter. Twitter and Facebook are followed by the professional networking platforms Xing and LinkedIn with regard to their relevance. In addition, blogs are used. However, forums and media-sharing platforms such as YouTube, Instagram and wikis are not perceived as particularly important and are only seldom used in the experts' everyday journalistic work.

Altogether, social media enable a fast and highly efficient overview of current topics. Moreover, using social media, journalists can gain a competitive advantage by quickly identifying and picking up relevant broadcasts from social media. The notion that social media provide better access to original sources by linking these sources via Twitter or Wikipedia is seen as beneficial. In addition, social media are used to initiate contact with experts, originators of coverage, or other sources. Moreover, background information about sources is collected via social media such as LinkedIn, and sources can be connected to other persons in their social networks to verify the credibility of these sources.

Evaluation of the quality of social media. The overall journalistic quality of social media was difficult for many of the participants to evaluate. Social media were perceived as an additional source that has to be checked thoroughly. This notion is summarized in the following comparison by one expert:

'It is a huge marketplace, in the truest sense of the word, and there are high-quality market stands but there are also the big noisy brats who are only good at selling things without much behind it [...]. If I as a journalist use a network, I have to assess and double-check the sources before I spread the news. This is the normal way, but with social media, I am even more aware of dangers.' [E5]

| | Years of professional experience | Position, professional experience and social media experience |
|-----------|----------------------------------|--|
| Expert 1 | 9 | Investigative journalist for well-known nationwide newspapers and a political print magazine Academic instructor and coach for investigative journalism specializing in online inquiry |
| Expert 2 | 22 | Journalist for a local newspaper Blogger |
| Expert 3 | 10 | Journalist for a well-known online news provider |
| Expert 4 | 11 | Freelance journalist for nationwide radio stations and well-known nationwide newspapers Author of a book that addresses social networks and journalism |
| Expert 5 | 7 | Editor for new media for a nationwide TV station Blogger |
| Expert 6 | 16 | Editor for a nationwide TV station with work experience at a well-known nationwide newspaper and at an American TV station Columnist for different newspapers Holder of three journalistic awards, one for online journalism |
| Expert 7 | 3 | Editor at a nationwide print magazine Holder of one journalistic award Blogger |
| Expert 8 | 5 | Freelance journalist for several well-known nationwide newspapers and online news providers Blogger |
| Expert 9 | 30 | Editor for a local TV station Academic instructor in journalism |
| Expert 10 | 5 | Editor and social media manager for the online news sites of two well-known nationwide newspapers Blogger |
| Expert 11 | 7 | Freelance journalist for several well-known nationwide newspapers and online news providers |
| Expert 12 | 6 | Investigative journalist for a local newspaper |
| Expert 13 | 12 | Freelance journalist for several well-known nationwide newspapers and online news providers Investigative journalist Founder of an initiative to support journalistic inquiry |
| Expert 14 | 3 | Freelance journalist for a nationwide TV station and for a local newspaper concentrating on an online format Blogger |

Table 2. Description of the sample

Altogether, it was positively highlighted that through social media, an additional information source evolved that often provides more authentic and less biased information. An additional important advantage is the breadth of information that is generated via social media. Thus, experts emphasized that through social media, they have access to information that they would not have otherwise.

'The example of Syria, the videos you can find on YouTube are technically bad quality but show content you would otherwise have no access to.' [E13] Intrinsic quality, believability/credibility: The credibility or believability of social media was perceived as ambivalent. In many ways, the participants assessed the credibility of information according to the person publishing it and less according to the platform of a social medium. Social media that do not allow anonymity were therefore preferred.

'Well, if I visit a blog I do not know, for example, the first thing I do is to click on 'About Us' to see who is behind the blog.' [E9]

Twitter also enabled a short introduction of the person. Business networks such as Xing provide much information on a person and were used to verify a source. In addition, in one of the interviews, experienced professionals who are associated with a high quality and reputation, such as the German journalist Stefan Niggemeier, were mentioned. Additionally, the VroniPlag-Wiki, which is concerned with checking dissertations from prominent persons for plagiarism, was attributed with high credibility even though its participants are anonymous. In this case, experts from a university environment have been integrated and are considered sufficiently skilled.

'The quality is quite high. Especially there (VroniPlag-Wiki), the experts are actually researchers and work at universities [...].' [E10]

Intrinsic quality, accuracy. To evaluate the quality of social media from a journalistic perspective, the experts found it very important that the original source was disclosed to assess whether the information was correct and objectively described. Correspondingly, in this context, linking the information to the original source was emphasized as being important. In addition, difficulties in verifying the accuracy of pictures and videos arose. As an example, material from war zones was mentioned because it is particularly difficult to assess if it is from the past or in fact shows pictures that were taped live in the war zone. Table 3 displays some of the new options that social media provide. Hence, it is possible to verify material stemming from abroad without actually being on the scene. In particular, the combination of different platforms and programs can be a huge qualitative benefit for the process of inquiry.

Intrinsic quality, objectivity. The criterion of objectivity describes the impartiality of information or whether it was composed from a political motivation. For example, impartiality is doubted in the context of technology blogs, as displayed in Table 3. This criterion overlaps with the criterion of credibility. However, it is interesting that in some cases, emotional impartiality was not even requested because journalists often used social media to obtain an atmospheric picture and thereby sought to obtain emotional statements.

Contextual quality, timeliness/actuality and relevance. With regard to relevance and, more importantly, actuality, social media were for the most part positively evaluated; in particular, Twitter was described as a real-time medium with abundant and current content and information. It was highlighted that journalists often came across relevant topics that were only later discussed in established news formats. However, the actuality of social media was also considered critical, and some journalists described social media as sometimes 'too current.' Hence, not only true but also false news can be spread very rapidly via Twitter. In this context, the term 'actuality pressure' was mentioned, tempting journalists to spread information without verifying it beforehand.

'[...] this actuality pressure. You are tempted to publish things faster to be the first.' [E3]

Accessibility quality, establishment of interaction and contact. As a further criterion of social media quality for the process of inquiry, the possibility to establish interactions and contact with the source was identified by the participants. The experts evaluated this feature of social media as rather positive but wished for more possibilities to search for and contact users and experts, features that are missing from the picture-sharing platform Instagram, for example.

'[...] found many pictures of the Boston Marathon, but then there was the problem that I could not contact the users who uploaded the pictures. Although there is Instamessage, [...] only users who use the app can write each other. [...] if I want to talk to the person who uploaded the picture [...] (I have to) leave my email address in the comments section [...], but then it is public and not very practical.' [E11]

| Criterion | Positive example | Problem area | Relevance for... |
|--|---|--|---|
| Intrinsic quality: believability/credibility | 'You find very authentic material [...]' [E11] | 'The problem [...] is that a lot is written anonymously [...] so that you do not really know who edited the entry.' [E6] | Twitter: ++ Facebook: ++ YouTube: - Blogs: o Forums: -- |
| Intrinsic quality: accuracy | Not applicable | '[...] if they claim that a picture was taken at a certain place, I say, 'Go somewhere and take a picture with your smartphone,' [...]. Then, I look at streetview and check which buildings there are in this area and tell him 'move on for 100 meters, take another picture and activate the GPS.' Then, the GPS data are written on the picture, and I can then read and check if he in fact is where he claims to be.' [E1] | Twitter: + Facebook: o YouTube: - Blogs: o Forums: -- |
| Intrinsic quality: objectivity | 'In my opinion, on Twitter, things are reported more objectively than in the whole tabloid press.' [E8] | '[...] if you look at technology blogs [...], there are so many blogs that concentrate on Google, Facebook, Twitter or Apple [...] you have to be careful because those technology blogs often are more interested in describing products in a very attractive way and forgetting about problems or privacy aspects [...]' [E5] | Twitter: + Facebook: + YouTube: o Blogs: - Forums: - |
| Contextual quality: Actuality/timeliness and relevance | 'You always find [...] more material, also primarily there [...]' [E4] | '[...] it tempts news media to announce things very fast [...] without having thoroughly checked it beforehand.' [E3] | Twitter: ++ Facebook: ++ YouTube: + Blogs: o Forums: - |
| Accessibility quality: establishment of interaction and contact | '[...] I am looking for people who are well versed in a specific field [...] such experts (I think) often use social media.' [E1] | '[...] in forums it is even more difficult than in blogs to find the right source, i.e., a real name to contact that person.' [E14] | Twitter: ++ Facebook: ++ YouTube: o Blogs: o Forums: - |
| Scale ranging from 5: ++ = 'very strong, to 1: -- = 'very weak (Rating of authors) | | | |

Table 3. Classification of the journalistic quality of social media

Comparing Table 3 with the framework by Wang and Strong (1996), we find intrinsic quality to be the most important property of data from social media. However, intrinsic quality can also be considered a major impediment to the use of information from social media because credibility in particular is an issue. Not surprisingly from the contextual quality category, actuality or timeliness and relevance were named as major properties of social media. Accessibility quality can be obtained especially for social media that do not allow anonymous contributions or at least allow for making contact with possible informants. Representational quality plays no role in our research context.

Operational scenarios. As part of the interviews, different operational scenarios were identified. They are valuable for the conception and planning of the system but can also elucidate the problem definition. As an example, the following scenario is presented:

The freelance journalist H. investigated social media on the night of the ‘Kanzlerduell’ on German television between chancellor Angela Merkel and her contender Frank-Walter Steinmeier on 09/01/2013 to appraise the debate from the perspective of social media users. Her article was published in an online magazine on 09/02/2013 at 07:00 am. First, H. watched the TV coverage after the debate on the four major broadcasting stations until 11:30 pm. Afterwards, she began an inquiry in the discussion forums of the largest parties. At 01:00 am, she visited Twitter, focusing on tweets by politicians and opinion leaders as well as comments on those tweets. At 02:00 am, she browsed her preferred political blogs, Facebook and the forums of the broadcasting stations and watched excerpts of the debate on YouTube, including the attached comments. By 04:00 am, H. had assembled four pages of handwritten notes and a Word document including 10 pages of direct quotes from social media. She believed at the time that she had an idea of the way the debate was perceived on social media, although the second part of her inquiry had to be handled rather quickly. She defined the storyline for her report and began to write the article at 4:30 am. At 6:40 am, H. uploaded her 3.5-page article into the online editing system and reflected on the process of inquiry. Particularly under these conditions, she wished for more tools that enabled a faster, more user-friendly and comprehensive inquiry to help her concentrate on writing her articles.

2.3 Summary of the problem definition

From the experts’ view, social media is a hugely important part of journalistic inquiry. The participants in our study saw social media as something commonplace that is already integrated into their journalistic routines. Social media give an overview of current topics and enable the establishment of contacts as well as communication with experts and informants. In this context, Twitter and Facebook were perceived as particularly relevant. With regard to the evaluation of social media quality, it was found that social media provide an important benefit as well as increased quality if applied appropriately. In this context, the content of social media needs to be thoroughly checked. At times, experts complained about the lack of technical tools. The most important intrinsic and contextual quality criteria to evaluate media qualities such as actuality, objectivity, and credibility are also relevant in the context of the use of social media for journalistic inquiry. In addition, social media enable the establishment of further interactions and contact that positively influence the evaluation of quality (accessibility quality). However, the interviews also revealed different problematic areas. For example, at times, information is published without prior verification (actuality), and sometimes, it is difficult to identify the initiator of the information (credibility). Table 4 summarizes the major findings from the interviews and lists which requirements of the suggested information system are addressed.

| | Implications for requirement analysis |
|-----------------------|--|
| Intrinsic quality | R1: Establish media integration to be able to check the identity and believability of informants, the accuracy of the information from other sources, and objectivity by finding other sources. R3: Identify information sources to check for credibility. Use indirect indicators to evaluate credibility. R5: Use indirect indicators to check for accuracy of the information. R6: Check for plagiarism. |
| Contextual quality | R1: Establish media integration to be able to use relevant and actual information from multiple social media sources. R2: Categorize information by actuality (date and time). |
| Accessibility quality | R1: Establish media integration to be able to establish contact with informants via social media. R4: Search for contact details of sources. |

Table 4. Empirical results and requirements

Moreover, the expert statements indicated that they have only limited technological knowledge. Only a few technical possibilities for evaluating and assuring quality were mentioned by the experts. Hence, there is a demand for a technical system that helps journalists to assess the quality of social media for

the inquiry process by applying established quality criteria. For this purpose, there are many possibilities in social media. Reports on social media provide not only the text of a document for the classification of its quality but also links, comments, and comparable data. Those variables, such as the number of citations, the actuality, or the total number of articles of a source, can be used to define the requirements of a system and represent the quality of social media from the perspective of journalistic inquiry. Although there are individual ranking systems in the journalistic context (e.g., Diakopoulos et al., 2012) that arrange tweets based on followers, retweets or content and even that assess the credibility of eyewitness reports, to our knowledge, there is no system that uses the strength of the entirety of social media.

3 Suggestion: Requirements of an Information System

Based on our literature review, the results of our empirical study and the identified usage scenarios, the requirements of a system for a social media search that accounts for journalistic quality criteria are outlined in this section. The requirements are the first step in the actual development of a journalistic search engine for social media.

Requirement 1: Media integration. Our empirical results and the scenario indicate that nearly all social media offer actual and relevant information that could be beneficial for journalistic inquiry. Therefore, the information system needs to integrate a multitude of social media, especially Twitter, Facebook, political blogs, and media-sharing platforms such as YouTube and forums and including the commentaries posted by users.

To address this requirement, we intend to use an ontology that describes the characteristics of the different social media and that serves as a common framework for the multitude of social media. More specifically, we use the Multimedia Metadata Ontology (M3O) (Saathoff and Scherp, 2010) because of its formal basis and its proven suitability for integrating heterogeneous and dispersed data sources (Scherp et al., 2011). The M3O is a so-called core ontology and can be used for the description of metadata from social media. Applying a four-step integration process, we have already shown that various existing models of metadata formats for media content can be integrated into the core ontology M3O and used (Scherp et al., 2009). An additional important aspect is the triangulation of different, complementary and often contradictory information in social media. For this purpose, the core ontology M3O offers an approach to the reification of information. In summary, this reification approach allows the representation and the illustration of relationships of different and contradictory information on the same topic in one knowledge base. For more details, the literature on M3O should be consulted (Saathoff and Scherp, 2010).

Requirement 2: Actuality. The main competitive edge in journalism is actuality. An information system therefore has to order contributions from social media by date and time. Additionally, the different social media, e.g., a specific Twitter account, the Facebook profile of a politician, or the blog of a political activist can be categorized by actuality using measures such as articles per week or user comments per week.

To provide the up-to-date delivery of information from social media for journalistic inquiry, the system will continuously request updated information from a list of social media that are continuously added to the system. New social media sources that possibly arise with new topics or political events can be manually added by journalists via keywords to continuously extend the system.

Requirement 3: Credibility. The evaluation of social media sources in terms of credibility must start with the identification of the author. For this task, the use of data from private and professional networking sites such as Facebook or LinkedIn, details from the imprint or the search for other contributions of the same author are obvious. Alternatively, the evaluation of credibility can build on indirect indicators such as the number of visits and user comments or the number of external links and their strength.

It will also be possible for journalists to manually classify certain social media sources as credible. This process will add another facet to the evaluation of the credibility of the information system that current search engines cannot cover. To support journalists in the evaluation of the credibility of specific contributions or authors, the system will calculate multiple metrics automatically, such as those mentioned above. Furthermore, the connection and use of background knowledge from the Web of Data is planned to add facts for specific topics. The aim of the Web of Data approach is to connect different dispersed sources of data with the help of standardized protocols and description languages to make them accessible (Berners-Lee, 2006). One of the most prominent examples is DBpedia (Zhang and Nasraoui, 2009), a structured version of Wikipedia. However, other data sources, such as weather services, can be used to evaluate the credibility of social media content such as photographs.

Requirement 4: Establishment of contact. If an unambiguous identification of the source is possible, journalists should be provided with contact details, such as email addresses, account names in social networks, Skype names and/or telephone numbers, to make the personal verification of information or its clearance for inclusion in an article easier.

Contact details are classified along different social media; for example, on blogs, user comments are for the most part posted by registered users, which does not allow for direct contact, while on Twitter, direct contact is possible. The options to make contact are therefore presented according to the media and will be easy to select by clicking.

Requirement 5: Accuracy. To evaluate the accuracy of information from social media content, the system will provide different indirect measures, such as the number of visits, user comments, the number of external links and their strength, rankings of the contribution, likes on Facebook, and tweets or re-tweets.

Requirement 6: Objectivity. To check the independence of information from social media sources, a search for plagiarism will be incorporated into the system that checks for similarities with other texts from the Internet. Thereby, the system automatically checks if essential parts of the contribution were taken from another data source without a reference. This functionality can be provided by the use of specialized tools to detect plagiarism. Additionally, a connection to existing Web search engines such as Google is planned to detect obvious copies on the Web.

4 Conclusions

This paper has presented the results of a design science study based on a literature review and an interview study with journalism experts. We have shown that all social media can have merits in journalistic inquiry but that their quality differs (research question 1). Furthermore, we derived requirements from the interviews (see Table 4) for a system that allows professional journalists to conduct journalistic inquiry in social media without abandoning journalistic quality criteria (research question 2).

We are aware that our work comes with some limitations. Our expert interviews are only a first step in examining the use of social media, quality evaluations, and scenarios for the use of social media in journalistic practice. However, with the classification of social media by journalistic quality in Table 3, we presented a first taxonomy that needs to be validated in future empirical research to allow for a more detailed formulation of the requirements of an information system. This taxonomy built on established journalistic quality criteria categorized with the help of the framework by Wang and Strong (1996). Thereby, we were able to establish a connection between journalistic research and IS research.

The results of our interviews also showed how journalists use social media in their daily routine. The methods used by our top-notch experts can act as an example for well-conducted “data journalism” but should be incorporated into an information system. The actual development of such an information system that supports journalistic inquiry represents a challenging task. One possible approach of an information system is the development of a search engine that is able to query different and selectable social media whereby the search can be restricted by region and time. The results of the query should

be ordered by journalistic quality, and it should be possible to subscribe to a feed with the search settings.

We are aware that the current state of our design science project has limitations with regard to its generalizability, particularly because the sample has a positive bias towards social media for journalistic research. Nevertheless, we strongly believe that the results obtained from our high-class sample of top-notch journalists are of value for IS research, as we follow an approach to comprehensively understand the evaluation of social media content in the journalistic search and selection process. Thus, we believe that this design science project has merits for both theory and practice. From a theoretical point of view, knowledge on the quality of media and information is extended to social media. Overall, the project serves as an example for the cooperation between media management and information systems research. For media practice, the project has the potential to increase the quality of journalistic inquiry. In addition, we plan to implement and evaluate the proposed search system for social media as part of our future work.

However, the target group for such a system is comparatively small; therefore, a business model that builds on advertisements is not feasible. Thus, another sustainable business model has to be found that accounts for the initial development costs, operating costs and the costs of the further development of such a system.

References

- Appelberg, J., Johansson, E., Nygren, G. and P. Baranowski (2014). "Social media in the professional work of Polish, Russian and Swedish journalists." *Journal of Print and Media Technology Research* 3 (2), 107–118.
- Auer, S.; Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R. and Z. Ives (2007). "DBpedia: A nucleus for a web of open data." *Proceedings of the 6th international The semantic web and 2nd Asian conference on Asian semantic web conference*, 722–735.
- Berners-Lee, Tim (2006). *Linked Data*. <http://www.w3.org/DesignIssues/LinkedData.html> (visited on 11/25/2014).
- Bogart, L. (2004). "Reflections on content quality in newspapers." *Newspaper Research Journal* 25 (1), 40–53.
- Burgoon, J. K., Burgoon, M. and C. Atkin (1982). *The World of the Working Journalist*. New York: Newspaper Advertising Bureau.
- Carta (2010). *Horst Köhler: Ein Rücktritt unter Blog-Mitwirkung*. <http://carta.info/28218/horst-koehler-ein-ruecktritt-unter-blog-mitwirkung> (visited on 11/25/2014).
- Castillo, C., Mendoza, M. and B. Poblete (2011). "Information credibility on Twitter." *Proceedings of the 20th international conference on World wide web*, 675–684.
- Cottle, S. (2011). "Media and the arab uprisings of 2011: Research notes." *Journalism* 12 (5), 647–659.
- Diakopoulos, N., De Choudhury, M. and M. Naaman (2012). "Finding and assessing social media information sources in the context of journalism." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2451–2460.
- Eppler, M. and D. Wittig (2005). "Conceptualizing information quality: A review of information quality frameworks from the last ten years." *Proceedings of the 2000 Conference on Information Quality*, 83–96.
- Garbett, A.T., Comber, R., Egglestone, P., Glancy, M. and P. Olivier (2014). "Finding "real people": trust and diversity in the interface between professional and citizen journalists." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3015–3024.
- Gladney, G. A. (1990) "Newspaper excellence: How editors of small and large papers judge quality." *Newspaper Research Journal* 11 (2), 58–71.
- Gladney, G. A., Shapiro, I. and J. Castaldo (2007). "Online editors rate web news quality criteria." *Newspaper research journal* 28 (1), 55–69.
- Gulyas, A. 2013. "The influence of professional variables on journalists' uses and views of social media." *Digital Journalism* 1 (2), 270–285.
- Harcup, T. (2009). *Journalism: Principles and Practice*. Thousand Oaks, California: Sage Publications.
- Heravi, B.R., Morrison, D., Khare, P. and S. Marchand-Maillet (2014). "Where is the news breaking? Towards a location-based event detection framework for journalists." *MultiMedia Modeling. 20th Anniversary International Conference, MMM 2014, Dublin, Proceedings, Part II*, 192–204.
- Hermida, A. (2010). "Twittering the news: The emergence of ambient journalism." *Journalism Practice* 4 (3), 297–308.
- Kaplan A. M. and M. Haenlein (2010). "Users of the world, unite! The challenges and opportunities of Social Media." *Business Horizons* 53 (1), 59–68.
- Knight, S. and J. Burn (2005): "Developing a framework for assessing information quality on the World Wide Web". *Informing Science Journal* 8, 159–172.
- Kuechler, W. and V. Vaishnavi (2008). "On theory development in design science research: Anatomy of a research project." *European Journal of Information Systems* 17 (5), 1–23.
- Lennon, S. (2003). "Blogging journalists invite outsiders reporting in." *Nieman Reports* 57 (3), 76–9.

- Machill, M., Beiler, M. and M. Zenker (2008). *Journalistische Recherche im Internet. Bestandsaufnahme journalistischer Arbeitsweisen in Zeitungen, Hörfunk, Fernsehen und Online*. Vistas Verlag, Berlin.
- Merrill, J.C. (1968). *The Elite Press: Great Newspapers of the World*. New York: Pitman Publishing.
- Messner, M. and M. W. DiStaso (2008). "The source cycle. How traditional media and weblogs use each other as sources." *Journalism Studies* 9 (3), 447–463.
- Neuberger, C. (2012). "Journalismus im Internet aus Nutzersicht." *Media Perspektiven*, 40–55.
- Paulussen, S. and R. A. Harder (2014). "Social media references in newspapers." *Journalism Practice*, 8 (5), 542–551.
- Picard, R. G. (2000). "Measuring quality by journalistic activity." In: *Measuring Media Content, Quality, and Diversity: Approaches and Issues in Content Research*. Ed. by R. G. Picard, Turku, Finland: Media Economics, Content and Diversity Project and Media Group.
- Qazvinian, V., Rosengren, E., Radev, D. and Q. Mei (2011). "Rumor has it: Identifying misinformation in microblogs." *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 1589–1599.
- Saathoff, C. and A. Scherp (2010). "Unlocking the semantics of multimedia presentations in the web with the multimedia metadata ontology." *Proceedings of the 19th international conference on World wide web*, 831–840.
- Scherp, A., Saathoff, C., Franz, T. and S. Staab (2011). "Designing core ontologies". *Applied Ontology* 6 (3), 177–221.
- Scherp, A., Eißing, D. and C. Saathoff (2009). "A method for integrating multimedia metadata standards and metadata formats with the Multimedia Metadata Ontology." *International Journal of Semantic Computing* 6 (1), 25–50.
- Schiffes, S., Newman, N., Thurman, N., Corney, D., Göker, A. and C. Martin (2014). "Identifying and verifying news through Social Media." *Digital Journalism* 2 (3), 406–418.
- Sikdar, S., Byungkyu Kang, O'Donovan, J., Hollerer, T. and S. Adah, S (2013). "Understanding information credibility on Twitter." *Proceedings of the 2013 International Conference on Social Computing*, 19–24.
- Sweetser, K. D., Porter, L. V., Chung, D. S. and Eunseong Kim (2008). "Credibility and the use of blogs among professionals in the communication industry." *Journalism & Mass Communication Quarterly* 85 (1), 169–185.
- Wang, R. Y. and D. M. Strong (1996): "Beyond accuracy: What data quality means to data consumers." *Journal of Management Information Systems* 12 (4), 5–33.
- Weischenberg, S., Malik, M. and A. Scholl (2006). "Journalismus in Deutschland 2005. Zentrale Befunde der aktuellen Repräsentativbefragung deutscher Journalisten." *Media Perspektiven*, 346–361.
- Welker, M. (2007). "Journalisten als Weblognutzer: Wer sie sind, was sie denken." In: Rau, H (Ed.) *Zur Zukunft des Journalismus in der Mediengesellschaft*. Peter Lang, Frankfurt, 95–116.
- Zeller, F. and J. Wolling (2010). "Struktur- und Qualitätsanalyse publizistischer Onlineangebote." *Media Perspektiven*, 143–153.
- Zhang, Z. and O. Nasraoui (2009). "Profile-based focused crawling for social media-sharing websites." *EURASIP Journal on Image and Video Processing*, Article ID 856037, 1–13.
- Zubiaga, A., Ji, H. and K. Knight (2013). "Curating and contextualizing Twitter stories to assist with social newsgathering." *Proceedings of the 2013 International Conference on Intelligent User Interfaces*, 213–224.