The use of big data in corporate communication

Christian Wiencierz and Ulrike Röttger
Department of Communication, University of Münster, Münster, Germany

Abstract

Purpose – The purpose of this paper is to illustrate the current state of research on the significance of big data in and for corporate communication and to introduce a framework which provides specific connecting points for future research. This is achieved by summarizing and reviewing the insights provided by relevant articles in the most significant scholarly journals. The paper also investigates trends in the literature.

Design/methodology/approach – On the basis of a systematic literature review, 53 key articles from 2010 to 2015 were further analyzed.

Findings – The literature review illustrates the potentialities of big data for corporate communication, especially with regard to the field of marketing communication. It also reveals a dramatic lack of research in the fields of public relations and internal communication with respect to big data applications.

Research limitations/implications – The online databases used in this paper comprised of refereed scientific journals with the highest impact factor in the respective disciplines. Journals with a lower impact factor and books were not included in the search process for this thematic analysis.

Practical implications – This paper provides a conceptual framework that describes four phases of strategic big data usage in corporate communication. The results show how big data is able to highlight stakeholders' insights so that more effective communication strategies can be created.

Originality/value – This paper brings together previously disparate streams of work in the fields of communication science, marketing, and information systems with respect to big data applications in corporate communication. It represents the first attempt to undertake a systematic and comprehensive interdisciplinary overview of this kind.

Keywords Public relations, Marketing communications, Corporate communications, Big data, Systematic review

Paper type Literature review

Dealing with heterogeneous and steadily increasing amounts of data is one of the most important contemporary challenges faced by corporations. Measuring what matters and translating big data into decision making are future priorities for corporations (Loebbecke and Picot, 2015). Big data technologies and methods make it possible to have innovations that can enormously increase the competitiveness of corporations. For example, because big data provides new insights into the effects of strategic communication, corporations can use it to improve their corporate communication. It is because of such insights that big data also has great potential for communication science (Chen *et al.*, 2012; Parks, 2014).

The novelist William Gibson said: "The future is here, it's just not widely distributed yet!" (Fuguitt, 2015, p. i). It is beyond dispute that the future – in the form of big data – is already here (Mayer-Schönberger and Cukier, 2013). The question we are asking is whether big data has reached the stage where it is distributed with regard to corporate communication. The aim of this paper is, first, to analyze the significance of big data in and for the component spheres of corporate communication, namely, marketing communication, public relations (PR), and internal communication. Second, it introduces a framework about big data applications in corporate communication in order to deduce specific connection

points for future research. We understand big data as high volume, often unstructured and heterogeneous information assets generated in an appropriate speed with computer and storage systems to make data-enriched decisions. For a number of years, it has been possible to recognize the growing importance of big data from the rapidly rising number of (scientific) publications on this subject (Chen *et al.*, 2012). However, the literature lacks an overview of initial research activities concerned with big data in the context of corporate communication and an overview of the results of such initial research.

We conducted an interdisciplinary integrative literature review in order to bring together previously disparate streams of work in the fields of communication science, marketing, and information systems (IS) with respect to big data applications in corporate communication. In so doing, we set out to shed light on this phenomenon and its rapidly growing importance. To analyze the journal articles that were identified, we differentiated them with respect to the three component spheres of corporate communication (marketing communication, PR, and internal communication) as well as with respect to the fields of activity pertaining to corporate communication (e.g. customer communication, brand communication, and media relations). In addition, we developed a framework and described four phases of a systematic application process of big data in corporate communication.

Conceptualization of corporate communication and big data usage

For a sophisticated understanding of the potentialities and limits of big data applications in the context of corporate communication, it is necessary to consider the distinct component spheres of corporate communication separately. These are marketing communication, PR, and internal communication (Zerfass, 2008). Whereas marketing communication refers primarily to the market environment and supports both the sale and purchase of products and services, PR focuses its strategies on the socio-political environment (political-administrative and sociocultural publics). The aim of PR is to ensure scope for action and to legitimize concrete strategies through social integration. Internal communication is directed inward toward the corporate public; it aims to enable and support common, collaborative performance.

A stronger thematic focus can be used to differentiate big data usage with respect to the various typical fields of activity pertaining to corporate communication. For example, customer communication, brand communication, and product communication are typical fields of activity in marketing communication, whereas staff communication and management communication are fields of activity pertaining to internal communication. PR is multifaceted and can be differentiated with respect to diverse fields of activity, such as media relations, investor relations, or issues management. It is important to stress that the fields of activity are not clear-cut – for example, brand and product communication. Moreover, some fields of activity can be assigned to all of the distinct component spheres of corporate communication – for example, corporate identity, and communication controlling. CEO communication is part of internal communication as well as part of PR. Brand and product communication are fields of activity that primarily belong to marketing communication, but they also include PR tasks – for instance, when the communication is aimed at the image or reputation of the brand or product. Figure 2 in the Results section provides an overview of this conceptualization of corporate communication.

Moreover, the term big data has to be defined: compared to conventional databases, big data consist of very large, complex, and variable amounts of data (volume). Also, the degree of heterogeneity is a significant difference to conventional data. Big data is composed of data with diverse formats, structures, and semantics, e.g. text comments, videos, or data generated from wearables (variety). Concepts, technologies, and tools are needed for the fast

and systematic storage, administration, and analysis of the heterogeneous data, in order to enable the retrieval of information within seconds (velocity) (Beyer and Laney, 2012; Chen *et al.*, 2012; Gandomi and Haider, 2015). Besides these key characteristics, veracity is also essential for making data-enriched decisions (IBM, 2015). Only if the measured data are reliable can corporations make profound communication decisions on the basis of such data. We incorporate the foregoing explanation within our definition of big data as follows: Big data are high-volume, high-velocity, high-variety, and high-veracity information assets generated through computer and storage systems in a way that makes these assets manageable and usable for organizations and persons.

Finally, we present an ideal-typical conceptual framework that describes the process of how to make big data manageable and usable for the different component spheres of corporate communication (see Figure 1). In the second part of the paper, after having pointed out the research gaps in each of these component spheres, we provide specific suggestions based on this framework, how these research gaps can be systematically examined.

This model is based on Bunte and Krohn-Grimberghe's (2014) explanation of general applications of big data, which we have applied specifically to corporate communication. We describe four phases of strategic big data usage in corporate communication. In the first phase, the communication problem requiring a solution and the communication objectives have to be determined precisely, which is to be achieved by means of big data. In the second phase, it is necessary to specify which communication processes are to be measured (variety), how frequently and extensively they are to be measured (volume), and at what speed they are to be measured (velocity). Moreover, the reliability (veracity) of the measured data must be ensured. Big data is only beneficial to corporations if they gain knowledge from the data for the purpose of decision making (Beyer and Laney, 2012; Gandomi and Haider, 2015). Data mining refers to systematic, computer-based methods of analysis that are used to mine vast mountains of data for knowledge. Accordingly, in the third phase, big data analytic methods are used to process the quantities of data in order to identify regularities, patterns, and hidden relationships in the big data. This phase is broken down into four analysis processes: descriptive analysis, diagnostic analysis, predictive analysis, and prescriptive analysis. With regard to the communication processes that are to be measured by those responsible for corporate communication, it is crucial to clarify what was communicated (descriptive analysis) before it is possible to analyze why it was communicated (diagnostic analysis). It is then necessary to analyze what is likely to be
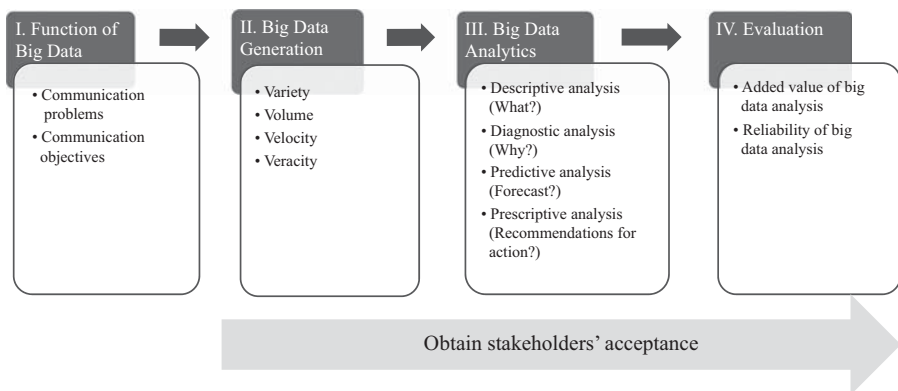
Figure 1.
Four phases of strategic big data usage in corporate communication

Source: Own depiction based on Bunte and Krohn-Grimberghe (2014)

communicated (predictive analysis) before recommendations can be made for action regarding corporate communication (prescriptive analysis) (Chen *et al.*, 2012; Gandomi and Haider, 2015; Mayer-Schönberger and Cukier, 2013). Finally, an evaluation of the big data applications should be conducted within the scope of the fourth and final phase.

Conduction of the systematic review of the use of big data in corporate communication

To provide an overview about big data applications associated with corporate communication we conducted a systematic, interdisciplinary literature review as proposed by Webster and Watson (2002). The review addresses the following research question:

RQ1. What significance does big data have in and for corporate communication?

We focused on reviewing articles with a high level of quality. Therefore, we considered impact rankings and journal ratings. In communication science, scholars do not frequently evaluate the quality of journals or provide rankings. For this reason, we took the impact ranking provided by Thomson Reuters into account and considered the seven journals with the highest impact factor. To meet our interdisciplinary requirement, we also considered high-quality articles from the fields of marketing and IS on the basis of journal rankings published by the German Academic Association for Business Research (VHB, 2015). In our review, we included the eight top-ranked journals in each marketing and IS – two of the sub-disciplines of business research. In addition, we included in our analysis the seven most internationally widespread communication science journals that deal with the subject of corporate communication and four journals focusing on big data. Our selection led us to examine a total of 34 journals (see Table I). We primarily used the EBSCOhost database to identify the relevant articles.

We defined search terms both from the domain of big data (big data, data analytics, data mining, business intelligence, and social media analytics) and from the domain of corporate communication (corporate communication, organizational communication, PR, brand communication, customer communication, communication management, strategic communication, marketing, advertising and advertisement, promotion, commercial, and publicity). We combined each of the terms associated with big data with each of the terms associated with corporate communication and searched within the key words and the full text of the article. The investigated time period was from January 1, 2010 to September 14, 2015. We did not consider articles that had been published before 2010 because we were interested in current developments. We read the abstracts of the articles we found. If it was obvious that the identified article did not even marginally address the significance of big data in and for corporate communication, the article was not included in the review. We examined the articles that became the basis of our review by reading the full papers. We identified 53 articles altogether (these titles are marked with an * in the reference list).

Results
Application of big data in corporate communication
The analysis of the articles that were considered indicates that the subject of most of the articles concerns the area of marketing communication. Only six of the papers either address PR as the subject of research or at least mention PR in some other way. Internal communication is the subject of research in only four articles (see Figure 2).

Marketing communication
Except for four articles, all the articles discuss potentials and challenges of big data for marketing communication. Some articles describe that big data will reshape business

Focus	Academic publication	Publications
Communication Science	*Journal of Communication*	x
	Journal of Computer-Mediated Communication	x
	Journal of Advertising Research	9
	Communication Research	x
	Management Communication Quarterly	x
	New Media & Society	1
	Human Communication Research	x
Marketing	*Journal of Marketing Research*	4
	Journal of Marketing	2
	Journal of Consumer Research	1
	Marketing Science	3
	Journal of Applied Psychology	x
	International Journal of Research in Marketing	3
	Journal of the Academy of Marketing Science	2
	Journal of Retailing	x
Information Systems	*Information Systems Research*	3
	Management Information Systems Quarterly	6
	Journal of Management Information Systems	3
	Mathematical Programming	x
	Journal of the Association for Information Systems	x
	Journal of Information Technology	x
	Information Systems Journal	x
	The Journal of Strategic Information Systems	4
Internationally widespread communication science journals with reference to corporate communication	*Journal of Advertising*	x
	International Journal of Strategic Communication	1
	Public Relations Review	2
	Journal of Public Relations Research	x
	Corporate Communications: An International Journal	2
	International Journal of Business Communication	x
	European Journal of Communication	x
Journals with a focus on big data	*Big Data and Society*	2
	Journal of Big Data	x
	Big Data Research	1
	Journal of Interactive Marketing	4

Table I.
Considered journals for the literature search

intelligence, that is, the collection and procession of data relevant to a corporation's market to gain insights supporting decision making, e.g. to improve marketing communication (e.g. Chi *et al.*, 2010; Esper *et al.*, 2010; Garg *et al.*, 2011; Hu *et al.*, 2014; Sahoo *et al.*, 2012). A particular focus is on describing the potential for increasing the chances of selling more products and services by improving brand communication (20 articles) and product communication (16 articles) through big data applications.

Especially advertising measures can be optimized by big data in order to maintain or build brand equity (Fulgoni, 2013; Rappaport, 2014; Seddon, 2015), introduce new products more effectively (Hariharan *et al.*, 2015), and improve customer acquisition through better targeting of advertising (Liu *et al.*, 2015) as well as translate insights into advertising effectiveness in conjunction with in-store marketing tactics (Harvey *et al.*, 2012). Martens and Provost (2014) present a big data approach that can be used to analyze the content of web pages. This approach enables web pages to be classified and allows advertisers to select which of these pages they want their advertisements to appear on.

In the context of a communicative brand and product profiling, the problem how to handle user-generated content (UGC) is mentioned in the articles. With big data technologies, dynamic analyses can be carried out to understand the effects of UGC in its

Internal Communication (4 articles)	Marketing Communication (49)	Public Relations (6)
Staff Communication (3)		Media Relations (0)
	Brand Communication (20)	
	Customer Communication (15)	
	Product Communication (16)	
		Investor Relations/ Finance Communication (1)
		Community Relations (0)
		Public Affairs/Lobbying (0)
		Issues Management (4)
		Crisis Communication (1)
		CSR Communication (2)
Management Communication/ CEO Communication (2)		CEO Communication (0)
	Corporate Identity (0)	
	Communication Controlling/Evaluation (28)	

Source: Own depiction based on Röttger *et al*. (2014, p. 190), Zerfass (2008, p. 80)

various forms (Goh *et al.*, 2013; Kilgour *et al.*, 2015). Campbell *et al.* (2011) present an approach for analyzing conversations about consumer-generated advertising. This refers to advertising produced by consumers themselves about brands they love or hate. Those responsible for managing advertising can use these approaches to understand feedback on consumer-generated advertising in the form of ad hoc comments and discussions on video-hosting sites. Tirunillai and Tellis (2014) provide communication professionals with a framework that they can use to observe product reviews concerned with competing brands, which enables managers to track in detail how this competition varies over time. According to Netzer *et al.* (2012), communication professionals can monitor the diffusion of information about brands and products in globally interconnected networks by using a combination of a text mining approach and semantic network analysis tools. They can also measure brand sentiments by monitoring word of mouth with social media analytics software (see also Schweidel and Moe, 2014).

In total, 15 articles describe the potentials to improve the customer communication, e.g. within the scope of customer relationship management (Malthouse *et al.*, 2013). In the field of customer communication, corporations can use big data to achieve the communication objective of personalizing and expanding service (Rust and Huang, 2014; Zablah *et al.*, 2012). In this way, big data results in deeper customer relationships. For instance, corporations can optimize their online user forums for technical support by monitoring and analyzing user comments and the behavior of problem solvers (Jabr *et al.*, 2014).

Most of the articles discussing application possibilities of big data in marketing communication imply more or less clearly that the potentials in this component sphere of corporate communication lie primarily in improving communication controlling and evaluation. In particular, the potential for evaluating and improving campaigns, such as advertising campaigns in the field of brand communication, is a matter that is often discussed (e.g. Campbell *et al.*, 2011; Cui *et al.*, 2015; Fan *et al.*, 2015; Seddon, 2015). For example, O'Regan *et al.* (2011) developed a big data solution recording the idiosyncrasies and characteristics of the relevant markets. The cellular provider MTS used it for advertising and media planning, among other applications.

PR
Significantly less articles discuss the diverse field of PR as compared to marketing communication. One of the six articles is by Esper *et al.* (2010), who focus more on value creation through knowledge management with respect to marketing communication, however. In their model, they only imply that PR strategy and tactics can also be improved by such knowledge management. They do not elaborate this point any further.

Four of these articles at least mention the possibilities of using big data to carry out issues management (Kent and Saffer, 2014; Colleoni, 2013; Chen *et al.*, 2012). Taking the oil industry as an example, Uldam (2016) illustrates that individual actors can easily generate topics in social media by themselves and create public attention for their protest. At the same time, she emphasizes the potentials for companies to collect intelligence about such individual actors by monitoring social media. This intelligence enables companies to conduct risk assessments or sentiment analyses, by which they win a more thorough picture of the discussed topics and the public opinion on them.

In addition, Uldam (2016) as well as Colleoni (2013) address the possibility of using big data to conduct and improve corporate social responsibility (CSR) activities. Colleoni (2013) describes the possibilities for investigating the efficiency of PR strategies in social media in order to achieve convergence between the CSR agendas of corporations and the social expectations of stakeholders. It is ultimately this convergence that increases corporate legitimacy. To do so, organizations have to use data mining techniques, such as topic and sentiment analysis.

Triantafillidou and Yannas (2016) as well as Kent and Saffer (2014) suggest that big, generated information assets can be used to manage PR more strategically by conducting research on the target publics, which relates to communication controlling and evaluation. Kent and Saffer (2014) stay rather vague, though, similarly to their determination that these technologies can improve PR with regard to crisis management. However, they critically observe that PR professionals do not exploit the full range of technical potentialities, e.g. the delivery of content via mobile devices.

Finally, Chen *et al.* (2012) bring up the topic of investor relations/finance communication in one sentence, by mentioning the opportunity to listen to the voice of investors. In contrast, none of the articles considers the potential of big data for the other fields of PR, namely, media relations, community relations, public affairs/lobbying, CEO communication, and corporate identity.

Internal communication
The four articles covering the topic of internal communication are about how corporations can or should provide an infrastructure for sharing relevant, complete, current, and accurate information among employees or executives (Micu *et al.*, 2011; Weinberg *et al.*, 2013). This relates to information gained from integrating and processing raw data from available sources (Popovič *et al.*, 2014). Furthermore, communication within a corporation will also change due to new digital communication possibilities, inter alia, because face-to-face interaction will decrease (Loebbecke and Picot, 2015).

Four phases of strategic big data usage in corporate communication
Beyond the results of the literature review just presented, another noticeable finding in the analysis is the fact that no holistic big data application process is described in the articles. Mostly, the authors mention or describe partial aspects of such a process. Consequently, there is a need for conceptual frameworks like ours presented at the beginning, which illustrates how to make big data manageable and usable for corporate communication. To demonstrate that our framework is generally applicable for corporate communication, aspects of the process mentioned in the articles are outlined exemplarily hereinafter. This general applicability enables us to use the framework as foundation to deduce new directions for research on the use of big data in corporate communication in a next step.

Phase 1: communication problem. According to the framework presented, it is necessary to determine the communication problems and objectives in the first phase of the big data process; the ability of big data to solve the problems and achieve the communication objectives must also be determined in this phase of the process. The literature review demonstrates a wide range of application possibilities to solve different communication problems employing big data. Mostly, these are located in the field of marketing communication, such as the improvement of communication controlling and evaluation or the increase of effectivity of brand, product, and customer communication.

Phase 2: big data generation. Researchers have measured a variety of communication processes; this concerns the second phase of the big data process. They measured blog entries about corporations and products (Chau and Xu, 2012; Gopaldas, 2014), product reviews (Decker and Trusov, 2010; Tirunillai and Tellis, 2014), consumer conversations about consumer-generated ads uploaded on YouTube (Campbell *et al.*, 2011), and user messages on Facebook brand fan pages (Xie and Lee, 2015).

Besides the question of which communication processes should be measured (variety), the volume of the collected processes is also critical for extracting meaningful information from the collected data set (Singh *et al.*, 2011). For example, Netzer *et al.* (2012) mined over 868,000 consumer messages with more than 6 million sentences about car brands and models for their case study. These efforts were aimed at demonstrating how the proposed text mining approach is able to track dynamics in the market structure using the real-time stream of data provided by consumer forums.

With respect to the characteristic of velocity, big data applications make it possible to conduct single point analyses (Campbell *et al.*, 2011) as well as dynamic real-time analyses (Netzer *et al.*, 2012). They also make it possible to carry out longitudinal analyses that cover the period of a campaign lasting several weeks (McDonnell Feit *et al.*, 2013) or even analyses that span several years (Tirunillai and Tellis, 2014; Young and Page, 2014).

With regard to data collection, a crucial point is the question of veracity – that is, the question of whether the data are reliable. Decker and Trusov (2010) discuss this aspect in the context of UGC. According to them, communication professionals ought to develop powerful filters for detecting fake reviews that have been written not by consumers but by professionals with an interest in manipulation. Another problem concerns the neglect of grammatical standards and rules. If UGC is of interest, misspellings, and grammatical inaccuracies should be considered (Hu *et al.*, 2014; Netzer *et al.*, 2012). One fundamental problem associated with collecting valid data concerns users' willingness to provide personal data. There is evidence that some users provide profile information that is more accurate than that provided by others because of different levels of privacy concerns and different levels of sensitivity to security threats (Norberg and Horne, 2014; Park *et al.*, 2012; Wakefield, 2013). However, a potential of big data is to categorize stakeholders by adoption of geodemographic and ethno-cultural taxonomies, without any use of questionnaires (Dalton and Thatcher, 2015; Webber *et al.*, 2015).

Phase 3: big data analytics. After collecting the data, communication professionals have to conduct big data analytics in the third phase of the big data process. Within the scope of an initial analysis, communication professionals have to describe their data (descriptive analysis: What happened? What have stakeholders communicated?). The authors of the articles often define their own attributes or characteristics or consider existing theories and models in order to describe and classify the quantity of data. For example, Campbell *et al.* (2011) used Aaker's (1997) five brand personality dimensions (sincerity, excitement, competence, sophistication, and ruggedness) to classify their results.

Within the scope of a second analysis, communication professionals have to describe the reasons for the communication processes (diagnostic analysis: Why did it happen? What are the reasons for stakeholders' communication processes?). There are approaches that can help to explain the reasons for the communication processes that are measured. For example, in transnational corporations, ethnography can be used to help decipher the contradictions, dilemmas, and mysteries that underlie abstract models and patterns in order to give executives a unique means of understanding market realities (Cayla and Arnould, 2013).

Within the scope of a third analysis, communication professionals should make predictions about future data characteristics (predictive analysis: What will happen? What will stakeholders communicate?) before they can analyze recommendations for action regarding communication management within a fourth analysis (prescriptive analysis: What are the recommendations for management? What should the corporation communicate?). The analysis of the articles emphasizes that the predictive analysis, i.e. the creation of forecasts, is almost not discussed or examined, for instance by connecting past data with present data and probability models. Kent and Saffer (2014) mention the need for social media research to move into prediction. Apart from that, recommendations for action based on analyses of the current state are derivated in some articles. With reference to the example of the US automotive market, Du *et al.* (2015) found empirical evidence to support their assumption that search trends relating to product features, which are readily available in near real time, can be treated as reflective indicators of importance trends relating to product features. Managers can use this big data application to dynamically adjust their advertising spending on the basis of the market response model they have developed. In addition, Ho *et al.* (2010) describe a visual data mining tool for identifying core customers for representing the relationship between customer preferences and product and brand positioning in an understandable way by visualizing large marketing data sets on a map (see also Hariharan *et al.*, 2015; Tirunillai and Tellis, 2014). From this, targeted advertising measures can be deduced.

Phase 4: evaluation. In the fourth and final phase, the big data structures and applications that have been implemented should be evaluated. Only with such a permanent evaluation can corporate communication fulfill the requirements coming with the big data characteristic veracity. For example, the quality of the data should be constantly examined. Accordingly, corporations have to instate data quality assurance systems (e.g. Park *et al.*, 2012). Further mechanisms of evaluation described in the articles have already been introduced in this paper, such as the filter to exclude reviews not written by real customers, or the validation of personal data.

New directions for research on big data applications in corporate communication
The articles identified illustrate that research on the significance of big data for corporate communication is only in its infancy and this topic is not widely discussed in the leading journals – especially in communication science. Our literature review shows initial research in marketing communication and highlights the dramatic research gaps in the areas of PR and internal communication. Potentials of big data for issues management were only

mentioned in a few articles. Future research should approach the question, what data can be generated in what way, in order to identify and analyze issues. This analysis enables communicators to deduce instructions and to strategically influence them. Connected to this is the question of potential of big data for an early capture and analysis of first signals of possible problematic situations and dangers, which can be addressed within the scope of a crisis PR based on big data analyses. In the field of media relations, researchers have to examine what data and what analyses can contribute to an improvement of the communication to multiplicators in order to address them more strategically. While social media steadily gains relevance, this research – as well as research in the area of marketing communication – should also consider multiplicators in these media, e.g. by developing network analyses for the identification of such multiplicators. Algorithm based analytical methods for the identification of semantic structures, such as text mining methods, are already examined in the field of marketing communication. They can also be employed for the strategic relationship management in the range of the capital and finance market, which becomes increasingly fast moving, requiring quicker reactions and thus, ongoing situation analyses. The exemplarily demonstrated approaches for future research emphasize the large gaps in research in the field of PR. Kent and Saffer (2014, p. 575) conclude, that if PR as a discipline does not become aware of the potential of new technologies, PR professionals will find themselves far behind the curve: "We risk driving farther down the road of irrelevance than we are already headed if all we can bring to an organization is our willingness to produce Facebook posts or tweets for our clients."

With regard to internal communication, the question raised in the articles concerning (big data) information management within a corporation is to be examined more profoundly. Part of this is also the following question: if and which information assets corporate communication can generate efficiently in a decentralized way, or if the generation should be conducted in a central data lab, providing every department with their specific needs, respectively? The internal information management could certainly be also an important component in order to be able to execute the communication concept of a corporate identity more efficiently. In that case, e.g. different analyses could be conducted based on a well-structured, consistent self-conception of the corporation. Moreover, information can be allocated efficiently among employees, suppliers, etc.

The demonstration of the ideal-typical application process primarily emphasizes that regardless of the component spheres of corporate communication, all of the raised questions share the essential problem of how the generation of knowledge can be improved, and by means of what data and what analyses: the ability to benefit from big data depends on whether the field of corporate communication is able to convert this data into knowledge (Yang and Kang, 2015). Our framework, which emphasizes the importance of prescriptive and predictive analysis, is also an encouragement to advance research on corporate communication toward an engagement with prediction. If reliable predictions by means of big data analyses are possible, this significantly improves the basis of decision making among all of the component spheres of corporate communication. However, the first goal of these research activities on predictive analysis is to examine if such analyses can be used as content for big data applications in corporate communication at all. Predictive analyses based on big data can be applied in the industry, e.g. to anticipate the durability of machine parts in a smart factory. Nonetheless, communication is a lot more diffused than machine parts and, thus, significantly more challenging to predict. Future research should examine if and what predictive models can be developed for corporate communication, e.g. to face possible communicative cases of conflict with an adequate crisis PR.

Our systematization of corporate communication's fields of activity is one theoretical contribution of this paper. This systematization is applicable to demonstrate which specific areas of corporate communication are already explored and which are yet to be explored.

Another contribution is our conceptual framework about the ideal-typical generation and application of big data, which fills the buzzword big data in these fields of activity with life. This is necessary because only if the question is clarified what is meant by the term big data, the much-needed scientific discourse about its potentials and challenges in corporate communication can be pursued. In this process, our paper is further an encouragement to subsume the research on the application of large, heterogeneous databases in corporate communication under the term big data. This enables the comparability of the existing research in the field. Especially because the term and connected recent developments are currently strongly discussed, the research should be conducted under the term big data.

Implications

Up to now, the general problem faced by corporations has been a lack of advanced analysis techniques and a lack of a commonly adopted methodology for effectively analyzing the relevant data sets at a time when there is a continuing trend of rapid growth in the volume of available data (Micu *et al.*, 2011). The aforementioned articles point out the enormous potential of constantly developed big data applications for the diverse fields of corporate communication, however. Communication professionals can analyze and optimize their communication activities to maximize performance either after the completion of an activity or in real time, regardless of whether it is a single or multi-platform activity, a transnational activity, or different activities conducted by the corporation in particular countries (Fulgoni and Lipsman, 2014; Seddon, 2015). Such insights into user and consumer behavior would not be possible without big data analytics.

The examples of big data applications illustrated in the literature review suggest that big data will lead to a paradigm shift in corporate communication. Mathematical analysis and the knowledge required to perform mathematical models are significantly changing the field of action associated with corporate communication (Loebbecke and Picot, 2015). In such instances, corporate communication is to be carried out in cooperation with IT and analysis specialists. Corporate communication will need skilled personnel who are able to explore, digest, synthesize, and explain incoming insights (Chen *et al.*, 2012; Micu *et al.*, 2011).

Finally, it is important to emphasize that, concerning big data applications, corporations must always consider a possible resistance from outside because of stakeholders' growing skepticism concerning the use of big data. Such skepticism grows because of privacy concerns and worries about the loss of freedom and independence (Clemons and Wilson, 2015; Newell and Marabelli, 2015). In order to address these concerns as well as to face data misuse, the profession of corporate communication should discuss and formulate ethical guidelines for big data applications.

Limitations

We are aware that relevant literature concerning big data in general has been published in a far greater number of journals than the 34 journals we considered as well as in conference proceedings, book sections, and working papers. We also concede that the domain of big data has existed for longer than the period that we examined. Although some authors indicate that most articles on this subject have been published in the recent past, it is necessary to analyze a larger amount of literature as well as literature from further back in the past in order to get a complete picture of the use of big data in corporate communication. Other search terms than the ones used for this literature review can identify further articles. Furthermore, literature with the subject of corporate communication should be analyzed that does not use terms like big data or the other search terms we considered, but describes data generations and analyses according to our big data characteristics. We have not considered this literature in our review. By including this literature, a more comprehensive

picture can be gained beyond this thematic analysis; these further analyses can also take our described ideal-typical conceptual framework as a basis.

Because our literature review considered the most important international journals in the disciplines of communication science, marketing, and IS, and because with big data, data mining, and data analytics we used the pivotal search terms, we have been able to provide a substantial overview of the use of big data in corporate communication. It is clear from the literature review that these massive amounts of data have become crucial for corporations and the communication professionals they employ. By analyzing our sample of articles, we have been able to demonstrate that the future of corporate communication is here in the form of big data technologies and methods. As Boyd and Crawford (2012, p. 662) wrote: "The era of big data has already begun". Corporations and the scientific community have only just begun to use big data to improve corporate communication and to research such communication activities, respectively. Nevertheless, there are still uncertainties and obvious gaps in research in the fields of PR and internal communication with respect to big data. We can therefore say: "the future is here, it's just not widely distributed yet!"

References

Aaker, J.L. (1997), "Dimensions of brand personality", *Journal of Marketing Research*, Vol. 34 No. 3, pp. 347-356.

Beyer, M.A. and Laney, D. (2012), "The importance of 'big data': a definition", Gartner Research Report, Stamford, CT.

Boyd, D. and Crawford, K. (2012), "Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon", *Information, Communication & Society*, Vol. 15 No. 5, pp. 662-679.

Bunte, S. and Krohn-Grimberghe, A. (2014), "Was bringt big data? Begriffserklärung, Nutzen und umsetzung [what is the benefit of big data? Definition, benefit, and implementation]", *Zeitschrift Führung + Organisation: ZfO*, Vol. 83 No. 6, pp. 372-378.

*Campbell, C., Pitt, L.F., Parent, M. and Berthon, P. (2011), "Tracking back-talk in consumer-generated advertising", *Journal of Advertising Research*, Vol. 51 No. 1, pp. 224-238.

*Cayla, J. and Arnould, E. (2013), "Ethnographic stories for market learning", *Journal of Marketing*, Vol. 77 No. 4, pp. 1-16.

*Chau, M. and Xu, J. (2012), "Business intelligence in blogs: understanding consumer interactions and communities", *MIS Quarterly*, Vol. 36 No. 4, pp. 1189-1216.

*Chen, H., Chiang, R.H.L. and Storey, V.C. (2012), "Business intelligence and analytics: from big data to big impact", *MIS Quarterly*, Vol. 36 No. 4, pp. 1165-1188.

*Chi, L., Ravichandran, T. and Andrevski, G. (2010), "Information technology, network structure, and competitive action", *Information Systems Research*, Vol. 21 No. 3, pp. 543-570.

*Clemons, E.K. and Wilson, J.S. (2015), "Family preferences concerning online privacy, data mining, and targeted ads: regulatory implications", *Journal of Management Information Systems*, Vol. 32 No. 2, pp. 40-70.

*Colleoni, E. (2013), "CSR communication strategies for organizational legitimacy in social media", *Corporate Communications: An International Journal*, Vol. 18 No. 2, pp. 228-248.

*Cui, G., Wong, M.L. and Wan, X. (2015), "Targeting high value customers while under resource constraint: partial order constrained optimization with genetic algorithm", *Journal of Interactive Marketing*, Vol. 29, pp. 27-37.

*Dalton, C.M. and Thatcher, J. (2015), "Inflated granularity: spatial 'big data' and geodemographics", *Big Data & Society*, Vol. 2 No. 2, pp. 1-15.

*Decker, R. and Trusov, M. (2010), "Estimating aggregate consumer preferences from online product reviews", *International Journal of Research in Marketing*, Vol. 27 No. 4, pp. 293-307.

*Du, R.Y., Hu, Y. and Damangir, S. (2015), "Leveraging trends in online searches for product features in market response modeling", *Journal of Marketing*, Vol. 79 No. 1, pp. 29-43.

*Esper, T.L., Ellinger, A.E., Stank, T.P., Flint, D.J. and Moon, M. (2010), "Demand and supply integration: a conceptual framework of value creation through knowledge management", *Journal of the Academy of Marketing Science*, Vol. 38 No. 1, pp. 5-18.

*Fan, S., Lau, R.Y.K. and Zhao, J.L. (2015), "Demystifying big data analytics for business intelligence through the lens of marketing mix", *Big Data Research*, Vol. 2 No. 1, pp. 28-32.

Fuguitt, G. (2015), "The future is here", *Journal of Advertising Research*, Vol. 55 No. 1, p. i.

*Fulgoni, G. (2013), "Big data: friend or foe of digital advertising? Five ways marketers should use digital big data to their advantage", *Journal of Advertising Research*, Vol. 53 No. 4, pp. 372-376.

*Fulgoni, G. and Lipsman, A. (2014), "Digital game changers: how social media will help usher in the era of mobile and multi-platform campaign-effectiveness measurement", *Journal of Advertising Research*, Vol. 54 No. 1, pp. 11-16.

Gandomi, A. and Haider, M. (2015), "Beyond the hype: big data concepts, methods, and analytics", *International Journal of Information Management*, Vol. 35 No. 2, pp. 137-144.

*Garg, R., Smith, M.D. and Telang, R. (2011), "Measuring information diffusion in an online community", *Journal of Management Information Systems*, Vol. 28 No. 2, pp. 11-38.

*Goh, K.-Y., Heng, C.-S. and Lin, Z. (2013), "Social media brand community and consumer behavior: quantifying the relative impact of user- and marketer-generated content", *Information Systems Research*, Vol. 24 No. 1, pp. 88-107.

*Gopaldas, A. (2014), "Marketplace sentiments", *Journal of Consumer Research*, Vol. 41 No. 4, pp. 995-1014.

*Hariharan, V.G., Talukdar, D. and Kwon, C. (2015), "Optimal targeting of advertisement for new products with multiple consumer segments", *International Journal of Research in Marketing*, Vol. 32 No. 3, pp. 263-271.

*Harvey, B., Herbig, T., Keylock, M., Aggarwal, R. and Lerner, N. (2012), "Exploding the legend of TV advertising and price promotions", *Journal of Advertising Research*, Vol. 52 No. 3, pp. 339-345.

*Ho, Y., Chung, Y. and Lau, K. (2010), "Unfolding large-scale marketing data", *International Journal of Research in Marketing*, Vol. 27 No. 2, pp. 119-132.

*Hu, Y., Du, R.Y. and Damangir, S. (2014), "Decomposing the impact of advertising: augmenting sales with online search data", *Journal of Marketing Research*, Vol. 51 No. 3, pp. 300-319.

IBM (2015), "The four V's of big data", IBM Big Data & Analytics Hub, available at: www.ibmbigdatahub.com/sites/default/files/infographic_file/4-Vs-of-big-data.jpg (accessed August 6, 2015).

*Jabr, W., Mookerjee, R., Tan, Y. and Mookerjee, V.S. (2014), "Leveraging philanthropic behavior for customer support: the case of user support forums", *MIS Quarterly*, Vol. 38 No. 1, pp. 187-208.

*Kent, M.L. and Saffer, A.J. (2014), "A Delphi study of the future of new technology research in public relations", *Public Relations Review*, Vol. 40 No. 3, pp. 568-576.

*Kilgour, M., Sasser, S.L. and Larke, R. (2015), "The social media transformation process: curating content into strategy", *Corporate Communications: An International Journal*, Vol. 20 No. 3, pp. 326-343.

*Liu, H., Pancras, J. and Houtz, M. (2015), "Managing customer acquisition risk using co-operative databases", *Journal of Interactive Marketing*, Vol. 29, pp. 39-56.

*Loebbecke, C. and Picot, A. (2015), "Reflections on societal and business model transformation arising from digitization and big data analytics: a research agenda", *The Journal of Strategic Information Systems*, Vol. 24 No. 3, pp. 149-157.

*McDonnell Feit, E., Wang, P., Bradlow, E.T. and Fader, P.S. (2013), "Fusing aggregate and disaggregate data with an application to multiplatform media consumption", *Journal of Marketing Research*, Vol. 50 No. 3, pp. 348-364.

*Malthouse, E.C., Haenlein, M., Skiera, B., Wege, E. and Zhang, M. (2013), "Managing customer relationships in the social media era: introducing the social CRM house", *Journal of Interactive Marketing*, Vol. 27 No. 4, pp. 270-280.

*Martens, D. and Provost, F. (2014), "Explaining data-driven document classifications", *MIS Quarterly*, Vol. 38 No. 1, pp. 73-99.

Mayer-Schönberger, V. and Cukier, K. (2013), *Big Data. A Revolution that Will Transform How We Live, Work and Think*, John Murray, London.

*Micu, A.C., Dedeker, K., Lewis, I., Moran, R., Netzer, O., Plummer, J. and Rubinson, J. (2011), "Guest editorial: the shape of marketing research in 2021", *Journal of Advertising Research*, Vol. 51 No. 1, pp. 213-221.

*Netzer, O., Feldman, R., Goldenberg, J. and Fresko, M. (2012), "Mine your own business: market-structure surveillance through text mining", *Marketing Science*, Vol. 31 No. 3, pp. 521-543.

*Newell, S. and Marabelli, M. (2015), "Strategic opportunities (and challenges) of algorithmic decision-making: a call for action on the long-term societal effects of 'datification'", *The Journal of Strategic Information Systems*, Vol. 24 No. 1, pp. 3-14.

*Norberg, P. and Horne, D. (2014), "Coping with information requests in marketing exchanges: an examination of pre-post affective control and behavioral coping", *Journal of the Academy of Marketing Science*, Vol. 42 No. 4, pp. 415-429.

*O'Regan, M., Ashok, K., Maksimova, O. and Reshetin, O. (2011), "Optimizing market segmentation for a global mobile phone provider for both targeting and insight", *Journal of Advertising Research*, Vol. 51 No. 4, pp. 571-577.

*Park, S.-H., Huh, S.-Y., Oh, W. and Han, S.P. (2012), "A social network-based inference model for validating customer profile data", *MIS Quarterly*, Vol. 36 No. 4, pp. 1217-1237.

Parks, M.R. (2014), "Big data in communication research: its contents and discontents", *Journal of Communication*, Vol. 64 No. 2, pp. 355-360.

*Popovič, A., Hackney, R., Coelho, P.S. and Jaklič, J. (2014), "How information-sharing values influence the use of information systems: an investigation in the business intelligence systems context", *The Journal of Strategic Information Systems*, Vol. 23 No. 4, pp. 270-283.

*Rappaport, S.D. (2014), "Lessons learned from 197 metrics, 150 studies, and 12 essays: a field guide to digital metrics", *Journal of Advertising Research*, Vol. 54 No. 1, pp. 110-118.

Röttger, U., Preusse, J. and Schmitt, J. (2014), *Grundlagen der Public Relations [Basics of Public Relations]*, VS Verlag für Sozialwissenschaften, Wiesbaden.

*Rust, R.T. and Huang, M.-H. (2014), "The service revolution and the transformation of marketing science", *Marketing Science*, Vol. 33 No. 2, pp. 206-221.

*Sahoo, N., Singh, P.V. and Mukhopadhyay, T. (2012), "A hidden Markov model for collaborative filtering", *MIS Quarterly*, Vol. 36 No. 4, pp. 1329-1356.

*Schweidel, D.A. and Moe, W.W. (2014), "Listening in on social media: a joint model of sentiment and venue format choice", *Journal of Marketing Research*, Vol. 51 No. 4, pp. 387-402.

*Seddon, J. (2015), "The brand in the boardroom. How Ogilvy & Mather reinvented the marketing principles of brand valuation", *Journal of Advertising Research*, Vol. 55 No. 2, pp. 146-161.

*Singh, S.N., Hillmer, S. and Wang, Z. (2011), "Efficient methods for sampling responses from large-scale qualitative data", *Marketing Science*, Vol. 30 No. 3, pp. 532-549.

*Tirunillai, S. and Tellis, G.J. (2014), "Mining marketing meaning from online chatter: strategic brand analysis of big data using latent Dirichlet allocation", *Journal of Marketing Research*, Vol. 51 No. 4, pp. 463-479.

*Triantafillidou, A. and Yannas, P. (2014), "How public relations agencies in Greece respond to digital trends", *Public Relations Review*, Vol. 40 No. 5, pp. 815-817.

*Uldam, J. (2016), "Corporate management of visibility and the fantasy of the post-political: social media and surveillance", *New Media & Society*, Vol. 18 No. 2, pp. 201-219 (First published online July 1, 2014).

VHB (2015), "German academic association for business research. VHB-jourqual3", available at: http://vhbonline.org/en/service/jourqual/vhb-jourqual-3/♯c5968 (accessed August 8, 2015).

*Wakefield, R. (2013), "The influence of user affect in online information disclosure", *The Journal of Strategic Information Systems*, Vol. 22 No. 2, pp. 157-174.

*Webber, R.J., Butler, T. and Phillips, T. (2015), "Adoption of geodemographic and ethno-cultural taxonomies for analysing big data", *Big Data & Society*, Vol. 2 No. 1, pp. 1-16.

Webster, J. and Watson, R.T. (2002), "Analyzing the past to prepare for the future: writing a literature review", *MIS Quarterly*, Vol. 26 No. 2, pp. xiii-xxiii.

*Weinberg, B.D., de Ruyter, K., Dellarocas, C., Buck, M. and Keeling, D.I. (2013), "Destination social business: exploring an organization's journey with social media, collaborative community and expressive individuality", *Journal of Interactive Marketing*, Vol. 27 No. 4, pp. 299-310.

*Xie, K. and Lee, Y.-J. (2015), "Social media and brand purchase: quantifying the effects of exposures to earned and owned social media activities in a two-stage decision making model", *Journal of Management Information Systems*, Vol. 32 No. 2, pp. 204-238.

*Yang, K.C.C. and Kang, Y. (2015), "Exploring big data and privacy in strategic communication campaigns: a cross-cultural study of mobile social media users' daily experiences", *International Journal of Strategic Communication*, Vol. 9 No. 2, pp. 87-101.

*Young, C. and Page, A. (2014), "A model for predicting advertising quality as a key to driving sales growth", *Journal of Advertising Research*, Vol. 54 No. 4, pp. 393-397.

*Zablah, A.R., Bellenger, D.N., Straub, D.W. and Johnston, W.J. (2012), "Performance implications of CRM technology use: a multilevel field study of business customers and their providers in the telecommunications industry", *Information Systems Research*, Vol. 23 No. 2, pp. 418-435.

Zerfass, A. (2008), "Corporate communication revisited: integrating business strategy and strategic communication", in Zerfass, A., van Ruler, B. and Sriramesh, K. (Eds), *Public Relations Research*, VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 65-96.

Corresponding author
Christian Wiencierz can be contacted at: christian.wiencierz@uni-muenster.de