BIG DATA AND AUTOMATION

IN STRATEGIC COMMUNICATION

Final manuscript.

Published as:

Wiesenberg, M., Zerfass, A., & Moreno, A. (2017).
Big data and automation in strategic communication.
International Journal of Strategic Communication, 11(2), 95-114

ABSTRACT

Big data and automation pose huge challenges for strategic communication. This article contributes to the limited body of knowledge in the field by introducing the concepts, outlining opportunities and potential problems, and identifying current perceptions and applications in the professional field. A large-scale survey, with respondents from across more than 40 countries, explores the expertise of communication professionals, applications within communication departments and agencies, and consequences for the profession at large. The study shows a large gap between the perceived importance and current practices, a lack of competencies and ethical reflection, and a limited use of opportunities. The full potential of big data analytics and algorithms has not been leveraged until now, which calls for new initiatives in the practice and further research.

KEYWORDS

Big data, Public relations, Survey, Communication strategy/planning, Measurement/evaluation, Automation

INTRODUCTION

Today's societies are transformed by the massive amount of data collected by organizations, intermediaries, technology firms, and platform providers: "Data is the oil of the information economy" (Mayer-Schönberger & Cukier, 2013, p. 16). However, the discussion about big data—how to acquire and use data from various sources to inform decision-making and deliver better products or services—has only very recently entered the realm of strategic communication (Weiner & Kochar, 2016). There is little academic research in the field, and, where there is, it is mostly linked to marketing communications and not to public relations. Nevertheless, communication practitioners of all disciplines need to be aware of the opportunities and challenges for their organizations. Big data might change their jobs dramatically, as "digitization and big data analytics (…) impact employment amongst knowledge workers—just as automation did for manufacturing workers" (Loebbecke & Picot, 2015, p. 149).

Automation, however, is another key challenge for strategic communication. It is closely linked to the availability of data streams. Here the focus is on using them to feed algorithms for creating and delivering content, and not only for analytics or decision-making. The huge potentials have been shown in political communication, for example, in the data-driven election campaign by President Obama in 2012 (Hersh & Schaffner, 2013; Kreiss & Jasinski, 2016; Nickerson & Rogers, 2014) or in 2016 with the increased usage of bots and automation (Kollanyi, Howard, & Woolley, 2016). There is also a great potential for communication science at large (Parks, 2014), including strategic communication and public relations (Holtzhausen & Zerfass, 2015).

Therefore, the main goals of this article are to a) outline the concepts of big data and automation; b) discuss opportunities and challenges for strategic communication and public relations; and c) empirically identify perceptions and applications of big data and automation in

the professional field. The empirical study covers different levels (the communication profession, communication departments and agencies, and individual communication practitioners) and focuses on one continent: it employs a quantitative survey among 2,710 professionals in 43 European countries.

LITERATURE REVIEW

In order to reach the goals outlined above, it is necessary to gain a profound understanding of big data and automated communication. In a second step, this review focuses on the existing literature about big data and automation in strategic communication. The last section brings these two strands together, by outlining research gaps and deducing research questions for empirical study from them.

Big data, analytics, algorithms, and automation

The rapid growth of data is not a new phenomenon. According to the Oxford English Dictionary, the term “information explosion” was introduced in 1941 to describe a “rapid increase in the amount of information available”. Two decades later, Marron and de Maine (1967) drew attention to the rapid growth of data and presented the idea of “automatic data compression”. Looking at the tremendous growth of several major communications media from 1960 to 1977, de Sola Pool (1983) concluded that, in this period, “much of the growth in the flow of information was due to the growth in broadcasting” (p. 610). The need to grasp and handle such information flows was born. In 1990, Denning pointed out the importance of “Saving All the Bits” and predicted:

> Both Autoclass and the genetic memory show that it is possible to build machines that can recognize or predict patterns in data without knowing the meaning of the patterns. Such machines may eventually be fast enough to deal with large data streams in real time. By the end of the decade they may well be advanced enough to serve on space probes and space-borne instruments, where they can monitor streams that would be incomprehensible to us directly. (p. 405)

Due to the rapid evolution of storage systems, digital storage became more cost-effective for storing data than paper in the mid-1990s (Morris & Truskowski, 2003). In 1997, Cox and Ellsworth used the term *big data* when describing the problem that "data sets do not fit in main memory (in core), or when they do not fit even on local disk" (p. 235). From then on, big data was associated with the huge amount of data that could no longer be stored by normal machines. In order to manage massive amounts of stored data that pushed "traditional data management principles to their limits", Laney (2001) introduced a "more formalized" approach to data management by controlling data *volume, velocity,* and *variety*. A decade later, those "Three V's" have emerged as a common framework to describe big data. While *Volume* describes a massive amount of stored data that provides new insights which were previously not available, *Variety* refers to multiple data types from different sources (e.g. text, pictures, sound). *Velocity* illustrates streaming data flows (data in motion) and their constant processing (Chen, Chiang, & Storey, 2012; Gandomi & Haider, 2015; Kwon, Lee, & Shin, 2014). Later, *Veracity* was added as the fourth "V", which emphasizes the limited reliability (or uncertainty) of the data collected and stored (Gandomi & Haider, 2015, p. 139). Today, these "Four V's of big data" have been widely accepted to define the phenomenon (Buhl, Röglinger, Moser, & Heidemann, 2013; Gandomi & Haider, 2015; Zikopoulos et al., 2013). Therefore, we propose the following definition:

> Big data denotes huge volumes and streams of different forms of data from diverse internal and external sources and their constant processing.

While there is a large agreement on these aspects in the literature and among experts, Gandomi and Haider (2015, p. 138) pointed out that many executives fail to understand big data in its full complexity because of its rapid evolvement. It is unknown whether this applies to communication professionals as well.

When big data is constantly and systematically collected, stored, and, finally, analyzed by computer-based methods we will refer to it as *data mining.* Data mining "uncovers interesting patterns and relationships hidden in a large volume of raw data, and the results tapped out may help make valuable predictions or future observations in the real world" (Che, Safran, & Peng, 2013, p. 3). This approach is called *predictive analytics.* It aims to prescribe actions or even control actions (Gandomi & Haider, 2015; Mayer-Schönberger & Cukier, 2013). However, such analytics require other algorithms, which are able to categorize and predict based on regularities, patterns, and relationships, as well as networks that are not visible at first glance. Hence, it is crucial to understand what kind of data has been mined (e.g. sources and their limitations), and how the algorithms analyzed them. That is why *descriptive analytics* (What occurred?), as well as *diagnostic analytics* (Why it occurs?) (Banerjee, Bandyopadhyay, & Tathagata, 2013), are also relevant (cf. Yaqoob et al., 2016 for detailed analysis techniques).

LaValle, Lesser, Shockley, Hopkins and Kruschwitz (2011) researched the use of big data analytics in a study among more than 3,000 business executives, managers, and analysts, from organizations around the world. They concluded that analytics is mostly used in a descriptive way to justify actions. This denotes an *aspirational* use of analytics. Predictive analytics (What might occur?), on the other hand, allows managers to rely on insights from data mining when guiding day-to-day operations. This is an *experienced* use of analytics. Prescriptive analytics (If X occurs, what should we do about it?) can even guide future strategies—in this case, the management process has been *transformed* by analytics (LaValle et al., 2011; see Table 1).

- Insert Table 1 here -

Moreover, LaValle and colleagues (2011) identified several key obstacles that need to be cleared away before using the full potential of big data analytics. The main obstacle is a lack of understanding of how to leverage analytics for business value. This is often based on insufficient statistical skills and a missing ability to connect insights from data to organizational and societal problems. All of this is mainly influenced by education and training, but also by work experience, and might differ across professions and industries. Another major obstacle has been identified on the organizational level, arising from disputes about the ownership of data and corporate cultures that do not encourage sharing information (e.g. a lack of cooperation between departments). Sometimes there are also concerns about the quality of the data, the ability to get data, and the costs involved (Schroeck, Shockley, Smart, Romero-Morales, & Tufano, 2013). Other authors highlight ethical, privacy, and legal concerns regarding big data analytics (Buhl, Röglinger, Moser, & Heidemann, 2013; Newell & Marabelli, 2015; Nunan & Di Domenico, 2013; Place, 2015). From the perspective of information technology professionals, the major challenges of working with big data are data security and risk management, lack of budget and time to study big data, as well as a lack of IT staff (Cisco Systems, 2013).

Big data analytics are closely linked to digital communication technologies and datafication (Boyd & Crawford, 2012; van Dijck, 2014). Algorithms programmed for data mining and algorithmic decision-making influence individuals or organizations and their communications, e.g. through the habitualized use of search engines or personalized social media platforms. Algorithms and analytics also influence society at large. They are part of the co-construction of

meaning and shape reality (Couldry & Hepp, 2016). Couldry, Fotopoulou, and Dickens (2016) emphasize the full range of analytics:

> 'basic analytics' (the automated measurement and counting installed within the operation of digital platforms, and associated websites, apps and tools); adjustments made by actors themselves so as to incorporate basic analytics into their daily practice; and thirdly, the architecture (the underlying organization of data flows) that allows digital platforms, and the measuring processes associated with them, to become (through a 'front end' design) embedded in the actions of those who interact with such platforms. (p. 119)

Tufekci (2015) argues that algorithms are "actants", "in that they are computational agents that are not alive, but that act with agency in the world" (p. 207). Algorithmic tools dynamically filter, highlight, suppress, or play an editorial role in determining which content is shown and whether it is shown or not. Moreover, "Facebook's News Feed and other such algorithmic decision makers 'decide' whether a news article shared by one of its users is shown to other users or not" (Tufekci, 2015, p. 208). Focusing on the consequences of such 'algorithmic decision-making' and, therefore, automation, Newell and Marabelli (2015, p. 6) pose two main concerns. First, in terms of big data, algorithmic decision-making might become superior to human judgement-based decisions, and this might create unfair discriminations. Second, in terms of micro-targeting and more granular data, they argue that monitoring an individual's behavior poses ethical concerns. This debate is largely reflected in the field of computational or algorithmic journalism (Anderson, 2012; Borges-Rey, 2016; Diakopoulos, 2016; Dörr, 2015; Fairfield & Shtein, 2014; Graefe, Haim, Haarmann & Brosius, 2016; Guo, Vargo, Pan, Ding, & Ishwar, 2016; Lewis, 2014; Lewis & Westlund, 2014; Lokot & Diakopoulos, 2016; Parasie & Dagiral, 2013; van der Haak, Parks, & Castells, 2012).

This first overview demonstrates that big data and big data analytics combined with algorithms and automation will not only change individual and organizational communication, but

also the social world and the public sphere. Couldry and Hepp (2016) characterize this as "deep mediatization". Consequently, the question arises as to how scholars from strategic communication and public relations have reflected upon these topics until now.

Big data, analytics, algorithms, and automation in strategic communication

The debate about *big data* is relatively new within the broader domain of strategic communication. A systematic and interdisciplinary literature review identified 53 articles published between 2010 and 2015 in 34 journals in this area (BLINDED, 2016a). Four papers have been published in 2010, five papers in 2011, seven in 2012, eight in 2013, 15 in 2014, and 14 in 2015.

Most of this research has been published in the field of *marketing communications.* Many scholars argue that big data supports micro-targeting of customers and co-creation of products and information, which positively contributes to brand, product, and customer communication, and, ultimately, helps to sell more products or services (e.g., Banasiewicz, 2013; Couldry & Turow, 2014; Erevelles, Fukawa, & Swayne, 2015; Fulgoni, 2014; Micu et al., 2011; Tirunillai & Tellis, 2014). Others emphasize the tremendous opportunities of big data for communication evaluation, measurement, and control, especially in social and online media, or by using sensors and other data points (e.g., Campbell, Pitt, Parent, & Berthon, 2011; Netzer, Feldman, Goldenberg, & Fresko, 2012; Rogers & Sexton, 2012). Apart from these conceptual considerations, the literature review did not identify any empirical research that demonstrates the diffusion of big data analytics in the marketing domain.

The debate is less developed in the field of *public relations.* In his programmatic book "The Automation of Public Relations", Phillips (2015) uses different scenarios to describe how big data and automation might affect the PR profession in general. In the same vein, the white paper by

Weiner and Kochhar (2016) outlines, from a public relations perspective, how organizations can implement big data analytics. They refer to the fields of evaluation for measuring "PR's impact on any number of factors being measured throughout the enterprise", tactics "in light of the overall business impact", strategy development with its focus on "conformists" audience targeting, objectives settings that "go beyond standard communications objectives", and landscape analysis that "tells the communicator about the past and present environment in which they are operating" (p. 15); not to forget predictive analysis to forecast the future environment and prescriptive analysis that offers different scenarios for decision-making (see Table 1 above). However, they also discuss some inherent risks of big data analytics, for example, the lack of basic conceptions from research and statistics (such as the difference between correlation and causality). First, it is necessary to understand what problems might be solved by utilizing big data. Second, the function of big data should be determined related to those objectives. Third the requirements of the four V's and various types of analysis (see Table 1) should be discussed, along with the communication strategy (p. 17). Finally, this analytical process has to be evaluated to ensure that the problems have been solved and the objective(s) have been reached.

However, this management process does not take into account privacy concerns and the need to gain acceptance from stakeholders who are targeted by the analytics. Potential threats of privacy pose ethical and legal challenges for advertising and marketing professionals. These challenges are deeply intertwined with data mining, big data analytics, and algorithms (Couldry, Fotopoulou, & Dickens, 2016; Gandomi & Haider, 2015; Newell & Marabelli, 2015; Yang & Kang, 2015).

Transferring these thoughts to *strategic communication* in general, Holtzhausen (2016) asks communication professionals to get involved in decisions on how algorithms are formulated and targeted. Practitioners should be aware of their ethical responsibility because the "ultimate aim

[of strategic communication] is to maintain a healthy reputation for the communicative entity in the public sphere" (Holtzhausen, 2016, p. 10). Hence, Holtzhausen and Zerfass (2015) conclude that strategic communicators should "become familiar with big data, its benefits and its shortcoming" (p. 13). They emphasize the usage of big data for strategic communication in the same two areas discussed above in the marketing literature, but add a critical connotation. On the one hand, they demonstrate the usage of big data for targeting individuals and other stakeholders within the public and the private sphere. They outline potential threats for individual privacy and the public sphere, as well as difficulties in balancing the interests of private communication and transparent public communication (p. 7). On the other hand, they highlight the potential of big data for measuring strategic communication outcomes, with the critical undertone that "strategic communicators still have to come to grips with the real value of these metrics" (p. 13).

The use of big data is deeply linked to the application of *algorithms.* As more aspects of strategic communication become digitized, "an increasing number of activities and processes central to the … function require computational interaction" (Collister, 2015, p. 365). Collister (2015) demonstrates three scenarios where algorithms challenge the communication profession: First of all there are "disruptive forces" for effective communication. For example, Facebook's auto-moderation function censors the online discourse to a certain extent, as it prevents some posts or comments "from appearing on a brand's page by automatically 'holding' … [them] unpublished for approval or deletion" (p. 364). Moreover, practitioners are able to study and reverse-engineer algorithms, such as Google's PageRank algorithm. With the understanding on how Google indexes and presents information, they can plan and implement a communications strategy to "primarily create and disseminate content that is designed solely to interact with and generate a positive outcome in search results" (p. 364). Finally, as noted above, algorithms play more and more of an essential role in the management process for strategic communication, as they help practitioners

get to grips with their internal and external environment. Social media monitoring tools can help to identify and track issues and clients, and measure brand performance or social media influencers (see also Ampofo, Collister, O'Loughlin, & Chadwick, 2015). These tools all "rely on algorithmic or computational processes to gather, interpret and understand the wider social environment" (Collister, 2015, p. 365). This means that practitioners ultimately hand over sense-making and interpretation of the public sphere to algorithms that, in turn, are used to plan or predict communication strategies (Collister, 2015).

Collister (2015) argues that "the non-human agency of computation highlights the indeterminability inherent in algorithmic PR" by using such tools as "black boxes" (p. 366). He makes the same point as Holtzhausen (2016) when he suggests the need for technical training for future professionals to understand how these "black boxes" are working. This insight of the literature review leads to the assumption that, although more and more communication professionals use such "black boxes", due to a lack of technical skills and knowledge, they probably do not really know about their functions and consequences.

This is even more relevant when algorithms are not only used for planning purposes, but when they trigger messaging routines, adapt content, or even create content automatically (Heimbach, Kostyra, & Hinz, 2015; Hoy, 2015; Lokot & Diakopoulos, 2016; Phillips, 2015). This is in line with current understandings of content strategies that include "all aspects of an organization's move to content management, from defining business goals and accounting for an organization's content to developing a company-wide strategy for producing, evaluating, governing, and publishing that content" (Andersen, 2014, p. 6). This poses opportunities and challenges for strategic communication as a discipline and practice, as a function in organizations, and as a profession where many practitioners will react to these developments with mixed feelings.

Adoption of big data by communication departments, agencies and professionals

The first part of this literature review offered an overview on the advent of big data, predictive analytics, algorithms, and automation and potential impacts on strategic communication. The following discussion sheds a light on the processes of adopting new technological options like big data on the organizational and individual level. Two theoretical approaches are especially helpful to explain these processes: New Institutionalism and the Extended Technology Adoption Model.

Within sociology, New Institutionalism (NI) comprises a set of middle range theories focusing on the meso level of organizations and the organizational field. Key concepts are isomorphism and practices of diffusion as well as the circulation of ideas (Boxenbaum & Jonsson, 2008; DiMaggio & Powell, 1983; Sahlin & Wedlin, 2008). Why and how are organizations "picking up popular ideas and seeking to incorporate them into their formal structures"? (Sahlin & Wedlin, 2008, p. 220). NI scholars claim that what was "most clearly observed to diffuse in such a way were often what appeared to be fashionable management ideas" (Sahlin & Wedlin, 2008, p. 220). These management ideas (techniques and models) are often introduced by consultants and come in waves (various forms of evaluations, assessments and rankings, certifications and accreditation processes as well as evidence-based guidelines). This effect can be observed in the field of strategic communication when organizations try to gain legitimacy (Sandhu, 2009; Schmeltz & Kjeldsen, 2016; Wehmeier, 2006). It seems also true for big data and its management as demonstrated by Van den Driest, Sthanunathan and Weed (2016). They propose the idea of building an "Insights Engine" to deeply understand "your customer's needs and fulfilling them better than anyone else" (p. 66) with the help of big data and its analytics, based on algorithms and automation. Hence, big data analytics can be seen as practice that emerges in organizations and different departments fueled by the activities of communication consultants and

service providers. It can be expected that those advisers are more aware of the discussion and more familiar with the concepts at hand, and that they have institutionalized such practices to a larger extent.

On the other hand, the Extended Technology Adoption Model (TAM2) is a widely used theory to conceptualize the decision-making of professionals in organizational settings (Venkatesh & Davis, 2000). Its focus is on the micro level of individual users and their acceptance of information technologies. The original Technology Adoption Model (TAM) aims "to provide an explanation of the determinants of computer acceptance that is general, capable of explaining user behavior across a broad range of end-user computing technologies and user populations, while at the same time being both parsimonious and theoretically justified" (Davis et al., 1989, p. 985). Based on two studies with two different computer programs, Davis identified two main factors that influence the actual use of technologies: the "Perceived Usefulness" (PU), i.e. whether a technology might improve work performance, and the "Perceived ease-of-use" (PEOU), related to the perceived level of additional effort necessary for using the new technology. Moreover, Davis et al. (1989) demonstrated that the behavioral intention for using a technology is influenced by PU and the attitude toward the technology. The actual use of a new technology can be predicted by the behavioral intention (Davis et al., 1989, pp. 985-989).

However, this model is situated in an ideal environment without any interference like time pressure, organizational or social influences. Therefore, the original TAM has been expanded by Venkatesh und Davis (2000). With the new model (TAM2), they demonstrated that both social influence (subjective norm, voluntariness, and image), perceptual processes (job relevance, output quality, result demonstrability, and PEOU), and experience significantly influence the acceptance of a new technology. Image in particular was identified as an important factor influencing technology acceptance on an individual level. This means that the personal experience of

communication professionals with big data, algorithms and automation, their familiarity with those concepts and their competencies in the field are relevant indicators for the use of these new technologies in strategic communication.

Research gaps and research questions

The literature reviewed showed that the knowledge about big data and automation in strategic communication and its subdomains is mainly shaped by conceptual discussions and case studies. This points to the need for empirical research. Scholars from different domains are convinced that the topics of big data, algorithms, and automation will change the communication profession, including the work of individual professionals and the work done by agencies and communication departments. Scholars have, on several occasions, emphasized the need for professionals to come to grips with big data. However, we do not know how communication professionals, the profession overall, and communication departments and agencies handle these issues. Based on these descriptive and theoretical insights, and in the absence of empirical data, the following research questions can be formulated:

RQ1: To what extent are communication professionals interested and skilled in the field and familiar with the concept of big data?

RQ2: To what extent and in which ways have communication departments and agencies implemented big data activities and algorithmic tools?

RQ3: How will big data transform the profession of strategic communication?

METHODOLOGY

The research is based on a quantitative survey among communication practitioners in Europe (BLINDED, 2016b). The questionnaire included a special section about big data and

automation which covered six questions derived from the literature review above. A pretest was held with 40 practitioners in 15 European countries. All recommendations have been taken into account and were used to finalize the survey instrument. The online questionnaire was made available throughout March 2016. More than 40,000 personal invitations were sent to professionals working in communication departments of all kind of organization and in communication agencies across all 50 European countries[1] via e-mail, with three reminders. Additionally, national branch associations and networks were asked to invite respondents. In total, 3,287 respondents completed the questionnaire in full. For the analysis, respondents who were not clearly identified as part of the population, such as scholars and students or practitioners from other regions, were excluded. This resulted in 2,710 responses, which were used for the study at hand. 58.1 per cent of the respondents were female (n = 1,574), and 41.9 per cent were male (n = 1,136). The majority of the respondents had more than 10 years of experience on the job (59.9 per cent, n = 1,622), followed by a group with 6–10 years of experience (22.9 per cent, n = 620), and one with less than 5 years of experience (17.3 per cent, n = 468). Most respondents worked in communication departments of either joint stock companies (19.5 per cent, n = 529), private companies (17.9 per cent, n = 486), governmental organizations (13.1 per cent, n = 355), or nonprofit organizations (11.9 per cent, n = 323). The rest worked as communication consultants (37.5 per cent, n = 1,017). Most respondents (28.0 per cent) came from Northern Europe (Scandinavia and the British Isles), followed by Central Europe (19.2 per cent), Southeastern Europe (17.6 per cent), Western Europe (14.5 per cent), Southern Europe (13.8 per cent), and Eastern Europe (6.8 per cent).[2]

[1] The selection of 50 countries is based on the official country list by the European Union (2016) and the Columbia Encyclopedia.

[2] These geographic regions were derived from the Columbia Encyclopedia.

To assess significant differences, the data are analyzed with SPSS, Version 22, using Pearson's chi-square (χ^2), independent samples T-tests and one-way ANOVA with post hoc Scheffé, as well as correlations using Pearson product-moment correlation coefficient (r), depending on the variable.

FINDINGS

The results of this study emphasize the potentials and challenges of big data and automation in the field of strategic communication. A majority of the professionals somehow have these topics on their agenda. However, only a minority possess a sound expertise in the field. The study reveals a huge gap between the necessity for communication professionals to deal with this evolution and the low skills sets and knowledge that exist today. However practitioners with a higher affinity towards social media tend to be more open to these developments. A similar picture emerged for the importance and implementation of practices for automated communication.

Big data interest, conceptualization, and skills of communication professionals (RQ1)

The literature review indicated that the debate about big data has broad implications for society today and that it influences strategic communication in multiple ways. Nevertheless, not all communication practitioners seem to be interested in this development. The data shows that three out of five communication professionals (59.3 per cent) follow the ongoing debate on big data, 44.4 per cent have given it attention, and 14.9 per cent have even given it close attention. The practitioners that report a (strong) interest mostly work for companies or agencies, while practitioners in nonprofit and governmental organizations are less attentive (see Table 2). There are also significant differences between professionals working at different hierarchical levels. Heads of communication and agency CEOs pay closer attention to the debate about big data (M =

3.54, SD = 1.09) compared to team or unit leaders (M = 3.47, SD = 1.15) and team members or consultants (M = 3.25, SD = 1.18; r = -0.097, N = 2,552, $p \leq 0.01$). A similar correlation exists between attention and age (r = 0.085, N = 2,710, $p \leq 0.01$) and attention and experience (r = 0.068, N = 2,552, $p \leq 0.01$). Older and more experienced communication professionals pay closer attention to the debate on big data.

- Insert Table 2 here -

In order to gain insights into their understanding of big data, the questionnaire showed respondents a list of eight different definitions of big data in a randomized order. Four of the items represented correct definitions derived from the four V's outlined above ("mass quantities of stored data that provide new insights which were previously not available" = Volume; "a variety of multiple data types from internal and external sources" = Variety; "a fast stream of data (data in motion) and their constant processing" = Velocity; "high and low quality data from trusted and untrusted sources" = Veracity). The other four items were related to the topic but did not represent the concept of big data ("customized creation of content for different stakeholders", "interpretation of relevant data for strategic decision making", "all kinds of information which is available in real-time", "a multitude of information from social media"). The respondents were asked to pick all appropriate definitions of big data. Only 0.9 per cent classified all eight items correctly (as either appropriate or wrong) and 6.2 per cent classified seven of eight correctly. This summed up to 7.1 per cent who can be categorized as highly familiar with the common understanding of big data. About the same portion of practitioners (7.4 per cent) mixed up almost everything and show a quite different understanding; while the majority is somehow or moderately familiar with the prevailing concept of big data (see Table 3).

- Insert Table 3 here –

Expertise is usually based on the familiarity with a subject and on the ability and willingness to reflect upon it. Thus, a hierarchical cluster analysis was performed based on the respondents' attention to the debate about big data (Q1), and their familiarity with the concept of big data (Q2), in order to identify groups of practitioners with different levels of big data expertise. A four-cluster solution without z-transformation (meaning Q2 tends to have a slightly bigger impact on cluster allocation) proved to be most reasonable. The largest cluster (54.7 per cent of the respondents, n = 1,483) includes practitioners who are interested in the debate and are familiar with the concept of big data ranging from some familiarity to close familiarity (four or more items correctly classified). They are labeled *Experts*. The second largest cluster (22.1 per cent, n = 599) also shows a reasonable understanding of big data but they have not paid much attention to the debate about the subject. Therefore, they were named *Informed.* Practitioners from the third identified cluster (17.0 per cent; n = 462) have paid (close) attention to the debate on big data but their familiarity with the concept is less or not at all developed. Hence, they are called *Bystanders*. The last identified cluster (6.1 per cent*;* n = 166) can be described as *Tenderfoots.* Those practitioners have neither paid attention to the debate nor developed a reasonable understanding of big data.

The Informed are less familiar with the concept of big data, however, they are also significantly less interested than the Experts when it comes to developing their skills and knowledge in the field. They also rate their understanding of the use of algorithms significantly lower than practitioners in all other clusters. Informed and Tenderfoots practitioners have significantly less social media expertise compared to the Experts and Bystanders (see Table 4).

Nevertheless, the results imply a connection between the self-perception of social media skills and an open-minded attitude towards big data. As mentioned before, only a minority really knows about the characteristics of big data. Moreover, the results reveal a backlog in technical skills and knowledge for all groups, especially regarding the understanding of algorithms and how to use them (e.g. by social media platforms or search engines) (see Table 4).

- Insert Table 4 here –

Interestingly, the Experts and the Bystanders clearly have more expertise, but they have also realized that they need even more technical skills and knowledge. The mean ratings are 3.36 (skills) and 3.53 (knowledge) in the top group of Experts practitioners, and it is even higher (M = 3.54 / M = 3.69 on a 5-point-scale) for the Bystanders. As such, the latter group reports the same level of deficiencies as the Tenderfoots, which is quite remarkable (see Table 4).

How is big data expertise among practitioners spread across different hierarchical levels? Table 5 reveals that the portion of Informed and Tenderfoots is almost the same in the circles of communication executives, unit leaders, and communicators working on the ground. Bystanders can be found significantly more often on the highest level—heads of communication departments and consultancies are aware of big data, but many of them do not know what it is really about. At the same time, the Experts are also strongly represented in the top ranks (41.0 per cent of this group); whereas only 24.8 per cent of this group are team members or consultants.

- Insert Table 5 here –

Big data and automation in communication departments and agencies (RQ2)

21.2 per cent of the respondents declared that their organization has implemented big data activities in the communication field, based on a definition of big data that included the four V dimensions mentioned above. The questionnaire showed this explicit definition to avoid a bias induced by knowledge deficits. 16.8 per cent stated that their department or agency plans to start such big data activities by the end of 2017, while 45.0 per cent indicated that their department or agency is not conducting such big data activities, and 7.6 per cent stated that they do not know how their organization handles the issue.

- Insert Table 6 here –

The analysis did not reveal statistically significant differences for the current grade of big data activities in different types of organizations (see Table 6). However, there are clear differences regarding the refusal of big data activities, which is quite strong in nonprofit organizations. The data also shows that a significantly higher portion of private and governmental organizations try to jump on the bandwagon by the end of 2017. It is not surprising that consultancies and agencies are in the lead when it comes to consulting others in the field of big data. Much more interesting is the finding that at least a quarter of communication departments consult internally about the topic. This demonstrates that strategic communication should not only comprehend big data as a tool for their own activities, but that it opens up new avenues for expanding its expert role in the organization.

- Insert Table 7 here –

Communication departments who want to excel in the field of big data need experienced practitioners, and communication professionals are most likely to gain experience or to show their skills and knowledge if they are working in an innovative organizational setting. It is thus

interesting to see how the different groups of practitioners identified in the last section are spread across the various types of organizations (see Table 7). Almost one quarter of the Experts (23.4 per cent), and even more of the Bystanders (29.2 per cent), work in organizations that have already implemented big data activities. At the same time, only 9.8 per cent of the Tenderfoots and 11.1 per cent of the Informed work in communication departments or agencies that are already active in the field. The highest level of expert practitioners (the Experts) can be found in communication agencies, but a reasonable part of this cluster is also employed in joint stock and private companies. The fact that most Bystanders (43.1 per cent) work in agencies might add to the prejudice that external consultants are good at spotting trends, but not always truly familiar with the concept of big data.

How are big data analytics used by communication departments and agencies? Those organizations in the sample who have already implemented big data activities (n = 508) use big data analytics most often to plan overall strategies (M = 3.77, SD = 1.00 on a 5-point Likert scale ranging from 1 = “Never” to 5 = “Always”), e.g. to use insights to guide future campaigns or for foresights (predictive and prescriptive). However, big data analytics is also frequently used to justify activities (e.g. by measuring results and demonstrating effectiveness) (M = 3.56, SD = 1.10). Analyzing big data to guide day-to-day actions (e.g. targeting publics with specialized content or through content adaptation) is also used quite frequently (M = 3.24, SD = 1.12). The results demonstrate that communication departments and agencies use the full potential of big data analytics, once they have begun to implement it. However, the predominant use for planning and evaluation reveals a rather traditional institutionalization, which does not leverage the full potential of big data. Utilizing big data for guiding day-to-day actions is a key innovation for strategic communication, which leads to automated operations like content distribution, adaptation, or even content creation.

The implementation of practices for automated communication was investigated through another instrument in the survey. All items were derived from the literature review. They represent different ways in which algorithms influence communication departments and agencies, and they show how algorithmic tools might be used in strategic communication (see Table 8). Less than one third of the respondents (29.2 per cent) indicated that their communication department or agency has implemented practices for adapting to algorithms of online services like search engines or social media platforms. Only 14.4 per cent of the organizations use algorithmic tools programmed to support decision-making (e.g. issues or crisis alerts, software to identify hot topics). A larger portion (23.6 per cent) has implemented algorithmic tools for fully or semi-automatic content distribution (e.g. for online channels, mailing lists, or newsletters). Algorithmic tools programmed for content adaptation, such as delivering targeted news on a corporate website, or content creation tools, for full or semi-automatically generating online content or press releases, are both used very rarely. Nevertheless, the study revealed a significant correlation between the usage of big data analytics to guide day-to-day actions and the adaptation to external algorithms and the implementation of algorithms for own activities (see Table 8). This means that big data and algorithms are indeed two sides of the same coin for strategic communication.

- Insert Table 8 here –

Transformation of the strategic communication profession through big data and automation (RQ3)

There is a relatively broad consensus that big data will change the communication profession (M = 3.85, SD = 0.83, *N* = 2,710; 5-point Likert scale ranging from 1 = “will not change at all” to 5 = “will substantially change”). In total, 72.3 per cent of the respondents agreed that big

data will change their profession. As a part of this group, 19.9 per cent even believe that big data will change the profession substantially. Only 6.8 per cent stated that big data will not change their profession or that it will not change their profession at all. Regarding the clusters, both the Experts (M = 4.07, SD = 0.75) and the Bystanders (M = 4.04, SD = 0.72) share a strong perspective that big data will change their profession, compared to similar, but less explicit, expectations of the Informed (M = 3.32, SD = 0.81) and the Tenderfoots (M = 3.22, SD = 0.80).

It is obvious that the strategic communication profession has not fully comprehended the effects of big data. This is also true for the understanding of algorithms and automated communication. Asking the respondents about the importance of various practices for automated communication for strategic communication today, they rated adapting to algorithms higher than all other items derived from the literature review (M = 4.03, SD = 0.98, *N* = 2,298; 5-point Likert scale ranging from 1 = “Not at all important” to 5 = “Extremely important”). However, algorithmic tools programmed to support decision-making (M = 3.83, SD = 1.13, *N* = 2,383), as well as algorithmic tools programmed for content distribution (M = 3.81, SD = 1.04, *N* = 2,345), are also important practices for strategic communication today. On the other hand, the relevance of algorithmic tools programmed for fully or semi-automatic content adaptation (M = 3.30, SD = 1.13, *N* = 2,384) or content creation (M = 3.27, SD = 1.23, *N* = 2,375) is supported by a majority of the professionals, but is clearly of lesser importance to them.

The results reported so far demonstrate the importance of big data and automation for the communication profession. On the other hand, a substantial gap between this importance and the level of implementation has been identified, both on the individual level of practitioners and on the organizational level. However, focusing on the profession in general, a final question asked respondents about the major challenges for communicators when working with big data. Very few (14.1 per cent) believe that ethical or legal concerns are one of the three most relevant issues when

working with big data in communications today (see Table 9). This has been frequently mentioned by scholars as an enduring problem for the field; however, professionals do not support this perspective or they have not thought about it to a great extent. The three most relevant challenges for the profession are the lack of analytical skills (mentioned by 48.6 per cent as a key issue), the lack of time to study or analyze big data (mentioned by 45.4 per cent), and the lack of technical skills. All of these challenges are related to the competencies of those working in the profession, rather than to structural or technical characteristics of big data and automation as such.

- Insert Table 9 here –

DISCUSSION

Organizations were never more in need of information expertise and information experts than in today's era of "information overload" (Micu et al., 2011, p. 218). Many professions, including information technology (IT), have profited from this development. Strategic communication and public relations, however, are, according to this large-scale study, lagging behind. Gaps on the individual, organizational, and professional level prevent strategic communication from becoming a fully developed insights and analytics function "that participates fully in business planning and organizational strategy" (Van den Driest, Sthanunathan, & Weed, 2016, p. 69). Instead, the trend, identified in a Delphi study by Kent and Saffer (2014), seems to be coming true. They wrote:

> If only one in ten of the predictions made in this study turn out to be true, public relations professionals will be far behind the ball when we look around and start asking why we do not know what is happening and have no advice to offer to our organizations or clients. 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 (p. 575).

For both scholars and practitioners of strategic communication, the social media environment has been primarily a source for listening or messaging towards stakeholders. This can partly be done by an "insight engine" and automated messaging (Van den Driest et al., 2016) based on algorithms. However, only one fifth of the communication departments and agencies in the sample have implemented such an insight engine until now, and a large part of the respondents does not even recognize the importance of such technologies. The potential of big data for communications is understood, but only at a quite abstract level. Specific knowledge and skills, as well as a sound implementation within organizations, are often missing. As long as practitioners have not got "to grips with the real value of these metrics" (Holtzhausen & Zerfass, 2015, p. 13) and keep focusing on operational activities and not on the strategic level (turning insights into strategy), they will continue losing ground against other functions in the organization.

The theoretical framework offers additional interpretations. Based on the empirical results of this large-scale study, big data can be seen as on overarching management fashion that has just reached the field of strategic communication. More and more communication departments and agencies try to jump on this bandwagon in the near future. New institutionalism suggests that implementing big data analytics and automation in strategic communication practice will continue to spread at least on the present rate. With regard to the Extended Technology Adoption Model (TAM2), the enormous gaps on the individual level might be traced back to the fact that the understanding and acceptance of these new technologies is missing. This interpretation is supported by the larger data set: professionals with higher social media expertise also report a higher expertise in the field of big data and automation. Based on TAM2, the low acceptance could be explained by a lack of understanding (communication professionals still do not know how these

new technologies might improve their daily work performance) and by a contradictory perception of the additional efforts for using such new technologies.

The establishment of big data and automation in organizations is strongly promoted by information technology and marketing experts, who are familiar with dealing with statistics and data, as demonstrated by Van den Driest et al. (2016). The data shows that consultants are ahead of communication departments in the field. From a theoretical standpoint, isomorphism, diffusion and the circulation of ideas (Boxenbaum & Jonsson, 2008; DiMaggio & Powell, 1983; Sahlin & Wedlin, 2008) explain why this drives the datafication of strategic communication. This might challenge the skills and education of practitioners. On the one hand, big data exacerbates the need to prove the impact of communication activities on organizational goals. More than thirty years ago, Grunig (1983) regretted the gap between the awareness of the need of evaluation and the actual failure of implementation of research and evaluation in public relations. Despite more than four decades of intense discussions, the issue of research and evaluation is still unsolved. Many organizations fail to evaluate communication in a way that links communications with organizational goals (Zerfass, Verčič, & Volk, 2017). However, on the other hand, the implementation of big data and automation brings unique opportunities for strategic communication. In order to design algorithms, a deep knowledge of stakeholder sets and the cognitions, attitudes, and behavior of key stakeholders is necessary. Knowledge about communication processes and opinion building is necessary to derive insights and create meaning from data.

Moreover, it is fundamental for strategic communication to expand the traditional goal of influencing stakeholders and to integrate a view from the outside in organizational decision-making. This can be done by interpreting the social and cultural context of communication

(Holtzhausen, 2016). According to the empirical results presented above, the communication profession tries to avoid falling further behind when they state that the major challenges are the lack of analytical skills (to make sense of big data) and the lack of time to study or analyze big data, as well as the lack of technical skills (to handle big data). Ethical and legal concerns are not seen as a major challenge. This might be related to knowledge deficiencies, and practitioners might also be aware of the reputational risk related to big data (Holtzhausen, 2016; Yang & Kang, 2015). However, closing the knowledge and skills gap and ethical and legal reasoning have to go hand in hand. Communication practitioners need to be well equipped with technical skills and knowledge about big data. They need to take part in building the algorithms that influence and even co-create reality in a world of (partly) automated communication. At the same time, they need to be aware of the potential risks inherent in big data, such as the threats to privacy (Boyd & Crawford, 2012; Child, Haridakis, & Petronio, 2012; Newell & Marabelli, 2015; van Dijck, 2014) or the discrimination or manipulation inherent in some algorithms (Boyd, Levy, & Marwick, 2014; Collister, 2015; Couldry, Fotopoulou, & Dickens, 2016; Diakopoulos, 2016; Tufekci, 2015; Woolley & Howard, 2016).

LIMITATIONS

The study provides an insight into the diffusion of big data and automation in strategic communication from a European perspective. However, it is important to underline that the survey gained a low response rate from Eastern Europe. Furthermore, the sample cannot be statistically representative of communication professionals in Europe, as the total population of practitioners working in the field of strategic communication is unknown. This study revealed the implementation of big data analytics and automation in communication departments and agencies;

however, the communication function is sometimes performed by other entities within an organization, including the executive board, and these could not be surveyed. Last but not least, the insights are based on reports by communication practitioners. The researchers were not able to cross-check the information by analyzing the actual performance of organizations or by testing the practitioner's expertise.

IMPLICATIONS AND PERSPECTIVES

This article reported about the first broader empirical research on big data, algorithms, and automation in the domain of strategic communication and public relations. It is a first step to building more knowledge in the field and opens many strands for further research. On the individual level of communication practitioners, the study revealed a low degree of familiarity with the concept of big data and limited skills in the field. This can be linked to future curricula in undergraduate and graduate courses, as well as to further training and education. Future practitioners in the field should not be data scientists, however, they need a comprehensive understanding of this highly sophisticated field. On the other hand, there is a huge research gap regarding the handling of communication tools driven by algorithms. The cluster analysis, demonstrated that a high number of practitioners use big data analytics and automated communication as a "black box" programmed by external agencies, without really having an understanding of how these algorithms work and what they do. Moreover, ethical and legal concerns about big data and how communication practitioners can defend the stakeholders' rights of privacy need to be tackled (Holtzhausen, 2016).

At the level of communication departments and agencies, this study revealed a low implementation rate of big data activities and algorithmic tools. The questions were quite broad to serve the exploratory character of this research. Further studies might reveal more by focusing on

analytical tools for big data communication used by organizations, and how organizations evaluate the performance of such tools. Different theories and concepts from New Institutionalism might guide these investigations. In addition, from a communication management perspective, the extent to which these activities and tools are guided by a sound management process needs to be investigated. Further research might also explore aspects of value creation: Has big data changed the way the communication function contributes to overall success? Have departments that use big data analytics established key performance indicators for this field? How does the implementation of big data activities affect the performance of communication departments and agencies? The view of top executives and other departments is as important as reports from communication professionals to shed light into these areas.

At the professional level, future research should focus on the concrete impacts of big data and automation on the identity and development of the field. A key question regards how and to what extent the communication profession will become more data driven. The same question arises for the topic of automation. First steps have been done by Collister (2015) for algorithmic communication, and by Holtzhausen (2016) for datafication. Comparative research will be very important as well—big data is a global phenomenon. Yang and Kang (2015) demonstrated that research across various cultures and countries is important in this field.

This broad study was not able to research the acceptance of big data analytics and automation by communication professionals in detail. However, the Extended Technology Adoption Model and results from this study regarding the level of experience, technical knowledge and skills might guide future further research on social influence and perceptual processes on the micro level.

The low expertise of the communication profession in the area of big data, algorithms, and automation revealed in this study is a cause for both concern and hope. On the one hand, the lack

of knowledge and skills needs to be critically reflected on by scholars and the profession alike. What does it mean to turn communication into data, data into insights, and insights into strategy? What is the role of algorithms in this process? Thus, scholars also need a comprehensive understanding to get to grips with algorithms and big data, in order to gain deeper insights into the impact of big data and algorithms used by strategic communication. On the other hand, the profession has just started to explore big data, as the study exposed. Scholars and the profession alike will surely delve deeper into the topic in the near future, and this can inspire innovation in multiple ways.

REFERENCES

Ampofo, L., Collister, S., O'Loughlin, B., & Chadwick, A. (2015). Text Mining and Social Media: When Quantitative Meets Qualitative, and Software Meets Humans. In P. Halfpenny & R. Procter (Eds.), *Innovations in digital research methods.* Los Angeles, CA: SAGE.

Andersen, R. (2014). The emergence of content strategy work and recommended resources. *Communication Design Quarterly Review*, *2*(4), 6–13. doi:10.1145/2721874.2721875

Anderson, C. W. (2012). Towards a sociology of computational and algorithmic journalism. *New Media & Society*, *0*(0), 1–17. doi:10.1177/1461444812465137

Banasiewicz, A. D. (2013). *Marketing database analytics: Transforming data for competitive advantage*. New York, NY: Routledge.

Banerjee, A., Bandyopadhyay, T., & Acharya, P. (2013). Data Analytics: Hyped Up Aspirations or True Potential? *Vikalpa*, *38*(4), 1–11.

BLINDED (2016a). Blinded for review.

BLINDED (2016b). Blinded for review.

Borges-Rey, E. (2016). Unravelling Data Journalism. *Journalism Practice*, *10*(7), 833–843. doi:10.1080/17512786.2016.1159921

Boyd, D., & Crawford, K. (2012). Critical questions for Big Data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662–679. doi:10.1080/1369118X.2012.678878

Boyd, D., Levy, K., & Marwick, A. (2014). The Networked Nature of Algorithmic Discrimination. In S. P. Gangadharan (Ed.), *Data and Discrimination: Collected Essays* (pp. 53–57). Washington, DC: New America.

Boxenbaum, E., & Jonsson, S. (2008). Isomorphism, diffusion and decoupling. In R. Greenwood, C. Oliver, R. Suddaby, & K. Sahlin (Eds.), *Handbook of organizational institutionalism* (pp. 78–98). London, UK: Sage.

Buhl, H. U., Röglinger, M., Moser, F., & Heidemann, J. (2013). Big Data: A Fashionable Topic with(out) Sustainable Relevance for Research and Practice? *Business & Information Systems Engineering*, *5*(2), 65–69. doi:10.1007/s12599-013-0249-5

Campbell, C., Pitt, L. F., Parent, M., & Berthon, P. (2011). Tracking Back-Talk in Consumer-Generated Advertising. *Journal of Advertising Research*, *51*(1), 224–238. doi:10.2501/JAR-51-1-224-238

Che, D., Safran, M., & Peng, Z. (2013). From Big Data to Big Data Mining: Challenges, Issues, and Opportunities. In B. Hong, X. Meng, L. Chen, W. Winiwarter & W. Song (Eds.), *Database Systems for Advanced Applications: 18th International Conference, DASFAA 2013, International Workshops: BDMA, SNSM, SeCoP, Wuhan, China, April 22-25, 2013. Proceedings* (pp. 1–15). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-40270-8_1

Chen, H., Chiang, R. H. L., Storey, V. C. (2012). Business intelligence and analytics: from big data to big impact. *MIS Quarterly*, *36*(4), 1165–1188.

Child, J. T., Haridakis, P. M., & Petronio, S. (2012). Blogging privacy rule orientations, privacy management, and content deletion practices: The variability of online privacy management activity at different stages of social media use. *Computers in Human Behavior*, 28(5), 1859–1872. doi:10.1016/j.chb.2012.05.004

Cisco Systems. (2013). The Potential and Challenge of Data. Retrieved from http://www.cisco.com/c/dam/en/us/solutions/enterprise/connected-world-technology-report/Global-Data-CCWTR-Chapter3-Media-Briefing-Slides.pdf

Collister, S. (2015). Algorithmic Public Relations: Materiality, Technology and Power in a Post-hegemonic World. In J. L'Etang, D. McKie, N. Snow & J. Xifra (Eds.), *The Routledge Handbook of Critical Public Relations* (pp. 360–371). New York, NY: Taylor & Francis.

Couldry, N., Fotopoulou, A., & Dickens, L. (2016). Real social analytics: A contribution towards a phenomenology of a digital world. *The British Journal of Sociology, 67*(1), 118–137. doi:10.1111/1468-4446.12183

Couldry, N., & Hepp, A. (2016). *The Mediated Construction of Reality*. New York, NY: John Wiley & Sons.

Couldry, N., & Turow, J. (2014). Advertising, Big Data and the Clearance of the Public Realm: Marketers' New Approaches to the Content Subsidy. *International Journal of Communication, 8,* 1710–1726. Retrieved from http://ijoc.org/index.php/ijoc/article/view/2166/1161

Cox, M., & Ellsworth, D. (1997, October). *Application-controlled demand paging for out-of-core visualization*. Paper presented at the VIS97 IEEE Visualization '97 Conference, Phoenix, AZ. Retrieved from https://www.evl.uic.edu/cavern/rg/20040525_renambot/Viz/parallel_volviz/paging_outofcore_viz97.pdf

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 319–340. doi:10.2307/249008

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science, 35*(8), 982–1003.

Denning, P. J. (1990). Saving All the Bits. *American Scientist, 78,* 402–405.

Diakopoulos, N. (2016). Computational Journalism and the Emergence of News Platforms. In S. Eldridge II & B. Franklin (Eds.), *The Routledge Companion to Digital Journalism Studies.* New York, NY: Routledge.

DiMaggio, P. J., & Powell, W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review, 48,* 147–160.

Dörr, K. N. (2015). Mapping the field of Algorithmic Journalism. *Digital Journalism, 4*(6), 1–23. doi:10.1080/21670811.2015.1096748

Erevelles, S., Fukawa, N., & Swayne, L. (2015). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research,* 69(2), 897-904. doi:10.1016/j.jbusres.2015.07.001

Europe. (n.d.). In *Columbia Encyclopedia online*. Retrieved from http://www.encyclopedia.com/doc/1E1-Europe.html.

European Union. (2016). *About the EU: Countries.* Retrieved from http://europa.eu/about-eu/countries

Fairfield, J., & Shtein, H. (2014). Big Data, Big Problems: Emerging Issues in the Ethics of Data Science and Journalism. *Journal of Mass Media Ethics, 29*(1), 38–51. doi:10.1080/08900523.2014.863126

Fulgoni, G. (2014). Numbers, Please: Big Data: Friend or Foe of Digital Advertising? Five Ways Marketers Should Use Digital Big Data to Their Advantage. *Journal of Advertising Research*, *53*(4), 372-376. doi:10.2501/JAR-53-4-372-376

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. doi:10.1016/j.ijinfomgt.2014.10.007

Graefe, A., Haim, M., Haarmann, B., & Brosius, H.-B. (2016). Readers perception of computer-generated news: Credibility, expertise, and readability. *Journalism*. doi:10.1177/1464884916641269

Grunig, J. E. (1983). Basic research provides knowledge that makes evaluation possible. *Public Relations Quarterly*, *28*(3), 28–32.

Guo, L., Vargo, C. J., Pan, Z., Ding, W., & Ishwar, P. (2016). Big Social Data Analytics in Journalism and Mass Communication: Comparing Dictionary-Based Text Analysis and Unsupervised Topic Modeling. *Journalism & Mass Communication Quarterly*, *93*(2), 332–359. doi:10.1177/1077699016639231

Heimbach, I., Kostyra, D. S., & Hinz, O. (2015). Marketing Automation. *Business & Information Systems Engineering*, *57*(2), 129–133. doi:10.1007/s12599-015-0370-8

Hersh, E. D., & Schaffner, B. F. (2013). Targeted Campaign Appeals and the Value of Ambiguity. *The Journal of Politics*, *75*(2), 520–534. doi:10.1017/S0022381613000182

Holtzhausen, D. R. (2016). Datafication: threat or opportunity for communication in the public sphere? *Journal of Communication Management*, *20*(1), 21–36. doi:10.1108/JCOM-12-2014-0082

Holtzhausen, D. R., & Zerfass, A. (2015). Strategic Communication: Opportunities and Challenges of the Research Area. In D. R. Holtzhausen & A. Zerfass (Eds.), *The Routledge Handbook of Strategic Communication* (pp. 3–17). New York, NY: Routledge.

Hoy, M. B. (2015). If this then that: an introduction to automated task services. *Medical reference services quarterly*, *34*(1), 98–103. doi:10.1080/02763869.2015.986796

Information explosion. (n.d.). In *Oxford English Dictionary online*. Retrieved from http://www.oed.com/view/Entry/95568?rskey=Dcjq4x&result=1

Kent, M. L., & Saffer, A. J. (2014). A Delphi study of the future of new technology research in public relations. *Public Relations Review*, *40*(3), 568–576. doi:10.1016/j.pubrev.2014.02.008

Kreiss, D., & Jasinski, C. (2016). The Tech Industry Meets Presidential Politics: Explaining the Democratic Party's Technological Advantage in Electoral Campaigning, 2004–2012. *Political Communication*. (00), 1–19. doi:10.1080/10584609.2015.1121941

Kollanyi, B., Howard, P. N., & Woolley, S.C. (2016). *Bots and Automation over Twitter during the First U.S. Presidential Debate. Data Memo 1*. Oxford, UK: Project on Computational Propaganda. Retrieved from http://politicalbots.org/wp-content/uploads/2016/10/Data-Memo-First-Presidential-Debate.pdf

Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, *34*(3), 387–394. doi:10.1016/j.ijinfomgt.2014.02.002

Laney, D. (2001). *3D Data Management: Controlling Data Volume, Velocity, and Variety*. Retrieved from http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big Data, Analytics and the Path From Insights to Value. *MIT Sloan Management Review*, *52*(2), 20–31.

Lewis, S. C. (2014). Journalism In An Era Of Big Data. *Digital Journalism*, *3*(3), 321–330. doi:10.1080/21670811.2014.976399

Lewis, S. C., & Westlund, O. (2014). Big Data and Journalism. *Digital Journalism*, *3*(3), 447–466. doi:10.1080/21670811.2014.976418

Loebbecke, C., & 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*, *24*(3), 149–157.

Lokot, T., & Diakopoulos, N. (2016). News Bots. *Digital Journalism*, *4*(6), 682–699. doi:10.1080/21670811.2015.1081822

Marron, B. A., & De Maine, P. A. D. (1967). Automatic data compression. *Communications of the ACM*, *10*(11), 711–715. doi:10.1145/363790.363813

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Boston, MA: Houghton Mifflin Harcourt.

Micu, A. C., Dedeker, K., Lewis, I., Moran, R., Netzer, O., Plummer, J., & Rubinson, J. (2011). The Shape of Marketing Research in 2021. *Journal of Advertising Research*, *51*(1), 213–221. doi:10.2501/JAR-51-1-213-221

Morris, R. J., & Truskowski, B. J. (2003). The evolution of storage systems. *IBM Systems Journal*, *42*, 205–217.

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine Your Own Business: Market-Structure Surveillance Through Text Mining. *Marketing Science*, *31*(3), 521–543. doi:10.1287/mksc.1120.0713

Newell, S., & 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*, *24*(1), 3–14. doi:10.1016/j.jsis.2015.02.001

Nickerson, D., & Rogers, T. (2014). Political Campaigns and Big Data. *Journal of Economic Perspectives*, *28*(2), 51–74.

Nunan, D., & Di Domenico, M. (2013). Market research and the ethics of big data. *International Journal of Market Research*, *55*(4), 505–520. doi:10.2501/IJMR-2013-015

Parasie, S., & Dagiral, E. (2013). Data-driven journalism and the public good: "Computer-assisted-reporters" and "programmer-journalists" in Chicago. *New Media & Society*, *15*(6), 853–871. doi:10.1177/1461444812463345

Parks, M. R. (2014). Big Data in Communication Research: Its Contents and Discontents. *Journal of Communication*, *64*(2), 355–360. doi:10.1111/jcom.12090

Phillips, D. (2015). *The Automation of Public Relations: A perspective on the development of automation affecting Public Relations*. Woolbridge, UK: BLURB.

Place, K. R. (2015). Exploring the Role of Ethics in Public Relations Program Evaluation. *Journal of Public Relations Research*, *27*(2), 118–135. doi:10.1080/1062726X.2014.976825

Rogers, D. & Sexton, D. (2012). Marketing ROI in the Era of Big Data: The 2012 BRITE/NYAMA Marketing in Transition Study. Retrieved from http://www8.gsb.columbia.edu/globalbrands/sites/globalbrands/files/images/2012-BRITE-NYAMA-Marketing-ROI-Study.pdf

Sahlin, K., & Wedlin, L. (2008). Circulating ideas: Imitation, translation and editing. In R. Greenwood, C. Oliver, K. Sahlin, & R. Suddaby (Eds.), *The SAGE handbook of organizational institutionalism* (pp. 218–242). Los Angeles, Calif.: Sage Publ.

Sandhu, S. (2009). Strategic Communication: An Institutional Perspective. *International Journal of Strategic Communication*, *3*(2), 72–92. doi:10.1080/15531180902805429

Schmeltz, L., & Kjeldsen, A. K. (2016). Naming as Strategic Communication: Understanding Corporate Name Change through an Integrative Framework Encompassing Branding, Identity and Institutional Theory. *International Journal of Strategic Communication*, *10*(4), 309–331. doi:10.1080/1553118X.2016.1179194

Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2013). *Analytics: The real-world use of big data. How innovative enterprises extract value from uncertain data.* Retrieved from http://www-935.ibm.com/services/multimedia/Analytics_The_real_world_use_of_big_data_in_Financial_services_Mai_2013.pdf

Sola Pool, I. de (1983). Tracking the flow of information. *Science*, *221*(4611), 609–613. doi:10.1126/science.221.4611.609

Tirunillai, S., & Tellis, G. J. (2014). Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation. *Journal of Marketing Research*, *51*(4), 463–479. doi:10.1509/jmr.12.0106

Tufekci, Z. (2015). Algorithmic harms beyond facebook and google: Emergent challenges of computational agency. *Colorado Technology Law Journal*, *13*(2), 203–217.

Van den Driest, F., Sthanunathan, S., & Weed, K. (2016). Building an Insights Engine. *Harvard Business Review*, *94*(9), 64–74.

Van der Haak, B., Parks, M., & Castells, M. (2012). The Future of Journalism: Networked Journalism. *International Journal of Communication*, *6*, 2923–2938.

Van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, *12*(2), 197–208.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*(2), 186–204.

Wehmeier, S. (2006). Dancers in the dark: The myth of rationality in public relations. *Public Relations Review*, *32*(3), 213–220. doi:10.1016/j.pubrev.2006.05.018

Weiner, M. & Kochhar, S. (2016). *Irreversible: The Public Relations Big Data Revolution (IPR White Paper)*. Gainesville, FL: Institute for Public Relations.

Woolley, S. C., & Howard, P. N. (2016). Automation, Algorithms, and Politics: Political Communication, Computational Propaganda, and Autonomous Agents. *International Journal of Communication*, *10*, 4882–4890.

Yang, K. C. C., & 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, *9*(2), 87–101. doi:10.1080/1553118X.2015.1008635

Yaqoob, I., Hashem, I. A. T., Gani, A., Mokhtar, S., Ahmed, E., Anuar, N. B., & Vasilakos, A. V. (2016). Big data: From beginning to future. *International Journal of Information Management*, *36*(6), 1231–1247. doi:10.1016/j.ijinfomgt.2016.07.009

Zerfass, A., Verčič, D., & Volk, S. C. (2017). Communication Evaluation and Measurement: Skills, Practices and Utilization in European Organizations. *Corporate Communications – An International Journal*, *22*(1). Doi: 10.1108/CCIJ-08-2016-0056

Zikopoulos, P., de Roos, D., Parasuraman, K., Deutsch, T., Corrigan, D., & Giles, J. (Eds.). (2013). *Harness the power of big data: The IBM big data platform*. New York: McGraw-Hill.

Table 1

Three stages of adopting analytics

	Aspirational	Experienced	Transformed
Motive	Use analytics to justify actions	Use analytics to guide actions	Use analytics to prescribe actions
Data management	Limited ability to capture, aggregate, analyze or share information and insights	Moderate ability to capture, aggregate and analyze data; limited ability to share information and insights	Strong ability to capture, aggregate and analyze data; effective at sharing information and insights
Analytics in action	Rarely use rigorous approaches to make decisions; limited use of insights to guide future strategies or day-to-day operations	Some use of rigorous approaches to make decisions; growing use of insights to guide future strategies, but still limited to use of insights to guide day-to-day operations	Most use rigorous approaches to make decisions; almost all use insights to guide future strategies, and most use insights to guide day-to-day operations
Type of analytics	Descriptive and diagnostic	From descriptive and diagnostic to predictive	Predictive and prescriptive

Note. Systematization based on LaValle et al., 2011, p. 24 (expanded).

Table 2

Attention to the debate about big data

	Companies	Governmental organizations	Nonprofit organizations	Consultancies & Agencies	Overall
Attention to the debate about big data *	3.44 (1.13)	3.15 (1.16)	3.31 (1.13)	3.55 (1.13)	3.43 (1.14)

Note. Standard deviations appear in parentheses bellow means. N = 2,710 communication practitioners in Europe. Q: Please rate these statements based on your experience. 5-point Likert scale ranging from 1 = "I have not given attention at all to the debate about big data" to 5 = "I have given close attention to the debate about big data". * $F(4,2705) = 9.229$. Highly significant at the $p \leq .01$level based on Scheffé post hoc test.

Table 3

Familiarity with the concept of big data among communication professionals

Familiarity with the concept of big data	*n*	%
Not familiar at all* (less than 3 items correctly classified)	200	7.4
Less familiar (3 items correctly classified)	428	15.8
Somehow familiar (4 items correctly classified)	640	23.6
Moderately familiar (5 items correctly classified)	833	30.7
Familiar (6 items correctly classified)	417	15.4
Very familiar (more than 6 items correctly classified)	192	7.1

Note. N = 2,710 communication practitioners in Europe. Q: "Big data" is characterized in various ways. Please pick all definitions which you believe are most appropriate. Big data refer to … * Including "None of these" / "I don't know" (n = 47).

Table 4

Big data expertise clusters among communication professionals (1/2)

	Experts	Informed	Bystanders	Tender-foots	Overall
I need to develop technical skills (program algorithms or websites; IT skills) *	3.36 (1.24)	3.31 (1.27)	3.54 (1.19)	3.54 (1.20)	3.39 (1.24)
I need to develop technical knowledge (understanding software algorithms, analytical understanding of big data, statistical knowledge) **	3.53 (1.20)	3.45 (1.25)	3.69 (1.13)	3.67 (1.29)	3.55 (1.21)
Understanding the use of algorithms (e.g. by social media platforms) ***	2.71 (1.11)	2.28 (1.02)	2.83 (1.13)	2.39 (1.15)	2.62 (1.11)
Social media skills (overall) ****	3.35 (.69)	3.05 (.75)	3.41 (.71)	3.02 (.84)	3.27 (.73)

Note. Standard deviations appear in parentheses bellow means. * N = 2,548 communication practitioners in Europe. Q: Thinking of yourself, your current capabilities and your future development, which of the following skills and knowledge areas do you believe are in need of developing? 5-point Likert scale ranging from 1 = "No need to develop" to 5 = "Strong need to develop". $F(3,2544) = 3.915$. Highly significant at the $p \leq .01$ level based on Scheffé post hoc test. ** N = 2,543. Same question and scale used. $F(3,2539) = 3.733$. Significant at the $p \leq .05$ level based on Scheffé post hoc test. *** N = 2,692. Q: How would you rate your personal capabilities in the following areas? 5-point Likert scale ranging from 1 = "Very low" to 5 = "Very high". $F(3,2688) = 31.090$. Highly significant at the $p \leq .01$level based on Scheffé post hoc test. ****N = 2,516. How would you rate your personal capabilities in the following areas? 5-point Likert scale ranging from 1 = "Very low" to 5 = "Very high", overall value based on a battery of 12 items. $F(3,2512) = 35.285$. Highly significant at the $p \leq .01$level based on Scheffé post hoc test.

Table 5

Big data expertise clusters among communication professionals (2/2)

	Experts %	Informed %	Bystanders %	Tender-foots %	Overall %
Head of communication department / Agency CEO	41.0	34.9	43.5	32.3	39.5
Team leader / Unit leader	34.2	32.5	34.9	35.4	34.1
Team member / Consultant	24.8	32.5	21.6	32.3	26.4

Note. N = 2,552 communication practitioners in Europe. Q: What is your position? Highly significant at the $p \leq .01$level based on Chi-square test, Cramér's V = .087.

Table 6

Big data activities in communication departments and agencies

My communication department / agency …	Joint stock companies %	Private companies %	Governmental organizations %	Nonprofit organizations %	Consultancies & Agencies %	Overall %
has implemented such big data activities	23.4	22.0	18.3	16.3	22.3	21.2
plans to start such big data activities until the end of 2017 *	17.7	21.1	18.3	16.0	14.1	16.8
is not conducting such big data activities **	43.1	42.7	49.7	54.9	42.1	45.0
consults (internal) clients and colleagues in the field of big data **	24.8	21.8	18.6	15.7	28.0	23.6

Note. N = 2,505 communication practitioners in Europe (excluding respondents who declared no knowledge about their organization's activities in the field by choosing the item "I don't know"). Q: "Big data" is mostly described as huge volumes and streams of different forms of data from diverse sources (external and internal) and their constant processing, which provide new insights. * Significant at the $p \leq .05$ level based on Chi-square test, Cramér's V = 0.069. ** Highly significant at the $p \leq .01$level based on Chi-square test, Cramér's V 'not conducting' = 0.089; Cramér's V 'consults' = 0.102.

Table 7

Big data expertise clusters in different types of organizations

	Experts %	Informed %	Bystanders %	Tender-foots %	Overall %
Joint stock companies *	19.6	20.0	20.6	13.9	19.5
Private companies *	18.2	17.4	17.1	19.9	17.9
Governmental organizations *	12.5	17.2	8.0	17.5	13.1
Nonprofit organizations *	11.4	12.7	11.3	15.7	11.9
Consultancies / Agencies *	38.2	32.7	43.1	33.1	37.5
	100	100	100	100	100
My communication department / agency has implemented such big data activities **	23.6	11.1	29.2	9.8	21.2

Note. * N = 2,710 communication practitioners in Europe. Q: Where do you work? Highly significant at the $p \leq .01$level based on Chi-square test, Cramér's V = .113. ** N = 2,505 communication practitioners in Europe. Q: "Big data" is mostly described as huge volumes and streams of different forms of data from diverse sources (external and internal) and their constant processing, which provide new insights. Taking into account this definition, my communication department/agency … Highly significant at the $p \leq .01$level based on Chi-square test, Cramér's V = 0.160.

Table 8

Implementation of practices for automated communication

	Percentage %	Overall *N*	Pearson* *r*
Adaptation to algorithms of online services like search engines or social media platforms **	29.2	2,435	.155
Algorithmic tools programmed to support decision-making	14.4	2,441	.054
Algorithmic tools programmed for fully or semiautomatic content distribution **	23.6	2,440	.175
Algorithmic tools programmed for fully or semiautomatic content adaptation **	7.0	2,417	.161
Algorithmic tools programmed for fully or semiautomatic content creation **	12.4	2,431	.208

Note. N^{min} = 2,431 communication practitioners in Europe. Q: Search engines and social media platforms use algorithms to select and display content. Similar approaches might be used by organizations to automate their communication activities. What is already used by your department/agency? * Correlation with "We analyze big data to guide day-to-day actions" (N = 508). * Highly significant at the $p \leq .01$level based on Pearson correlation.

Table 9

Major challenges when working with big data

Major challenges	%	*n*
Lack of analytical skills (to make sense of big data)	48.6	1,305
Lack of time to study/analyze big data	45.4	1,220
Lack of technical skills (to handle big data)	36.6	983
Data quality	31.1	837
Lack of software solutions fitting communication needs	25.5	686
Lack of budget	24.3	653
Organizational barriers (e.g. a lack of cooperation between departments)	22.8	613
Data security and risk management	22.1	595
Lack of IT staff who can support	15.8	424
Ethical concerns	14.1	379
Legal restrictions	13.6	366

Note. N = 2,687 communication practitioners in Europe. Q: In your opinion, what are the three (3) major challenges for the communication profession in general when working with big data? Percentages: Frequency based on selection as Top-3 challenge.