

Wanner, Jonas; Janiesch, Christian

Article

Big data analytics in sustainability reports: an analysis based on the perceived credibility of corporate published information

Business Research

Provided in Cooperation with:

VHB - Verband der Hochschullehrer für Betriebswirtschaft, German Academic Association of Business Research

Suggested Citation: Wanner, Jonas; Janiesch, Christian (2019) : Big data analytics in sustainability reports: an analysis based on the perceived credibility of corporate published information, Business Research, ISSN 2198-2627, Springer, Heidelberg, Vol. 12, Iss. 1, pp. 143-173, <https://doi.org/10.1007/s40685-019-0088-4>

This Version is available at:

<https://hdl.handle.net/10419/233160>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

Big data analytics in sustainability reports: an analysis based on the perceived credibility of corporate published information

Jonas Wanner¹  · Christian Janiesch¹

Received: 4 December 2017 / Accepted: 7 January 2019 / Published online: 14 January 2019
© The Author(s) 2019

Abstract The credibility of sustainability reports has been the subject of scientific research for several years. The problem is often referred to as the so-called credibility gap, which is based on information asymmetries. The situation is further complicated by the limited rationality of human action as improvements to reports do not necessarily translate into credibility gains. Research has proposed and extracted several methods to overcome the issue. Hitherto, most approaches to solve the problem focused on marketing-oriented approaches. This work takes a new approach and explores the extent to which information technology can increase credibility using the potential of big data analytics. We base our research on the relationship of the quality of information and on the perception of objective truth as postulated in the Habermas Theory of Communicative Action. We use the forecast-oriented Partial Least Squares Methodology for the review of hypotheses extracted from literature and expert surveys. The result confirms potential of the criteria of volume and veracity while velocity and variety do not yield comparable potential concerning sustainability reporting.

Keywords Big data analytics · Sustainability reporting · Credibility gap · Information quality

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s40685-019-0088-4>) contains supplementary material, which is available to authorized users.

✉ Jonas Wanner
jonas.wanner@uni-wuerzburg.de

Christian Janiesch
christian.janiesch@uni-wuerzburg.de

¹ Department of Information Management, Julius-Maximilians-University Würzburg, Josef-Stangl-Platz 2, 97070 Würzburg, Germany

1 Introduction

In recent years, more and more companies and organizations have been trying to make their activities more sustainable. To manage and measure own goals, performance, and operational changes, sustainability reports have become a very popular means (GRI 2016, 3). Sustainable actions and their transparency through reporting activities lead to several advantages perceived by external groups of interest such as an improved reputation (Weber 2014, 13f).

Despite the EU directive 2014/95/EU, the scope of this kind of reporting leaves companies with a lot of room to maneuver (Nachhaltigkeitskodex 2016). For example, criticism is directed at the traceability of the sources and the methods of data collection. The same applies to the completeness of measures and their information value (Knebel and Seele 2015, 198ff). This raises the question of the perceived credibility of published data and facts. The situation is regarded as precarious (Lock and Seele 2016, 186f; Nawratil 2006, 200ff) and is unfavorable for both sides—company and report recipient (Weber 2014, 97f). In science, the problem is referred to as the *credibility gap*, which is based on information asymmetries as a state of insecurity due to differing efforts and levels of information.

Regardless of initial successes, this gap has not been closed yet (Knebel and Seele 2015, 196; Lock and Seele 2016, 192). Previous approaches have in common that they only considered a situation where the discovery of information and the processing has already been completed. Until now, the focus has been chiefly on marketing-oriented approaches. We aim for a novel, alternative solution and investigate whether improvements to the credibility gap are possible before or during information retrieval and processing. Our approach considers big data analytics as the key factor for information generation.

The assumption of the potential improvement is based on the recognition that the perceived credibility correlates with the information quality of the report, which in turn depends on the quality of the data. This indication is supported by psychological credibility research. We examine whether an approach using innovative information processing capabilities, viz. big data analytics, can have a positive impact on data quality in sustainability reports. That is in particular: Does the perceived credibility of sustainability reports improve?

Summarizing, the starting point of our own investigation is the overlap between the properties of credible communication with those of information quality and its improvement by big data analytics. This leads to the following research questions:

RQ1: How can we conceptualize the credibility gap in sustainability reporting and operationalize information quality criteria to assess their impact on the perceived credibility of sustainability reports?

RQ2: Which criteria of information quality and which characteristics of big data have an influence on the credibility of sustainability reports and how can this finding be used to improve sustainability reporting?

In doing so, we aim to contribute to research on the (ir)rationality of decisions in business research and practice. We conceptualize criteria, which can guide complex management decisions regarding sustainability reports by examining the limited rationality of human action when consuming and perceiving reports, to derive practical recommendations on the effectiveness of big data analytics through its characteristics.

The paper is organized as follows: we begin with the foundations of sustainability reporting, perceived credibility, and the credibility of communication. The following section comprises the state-of-the-art of scientific research on the credibility gap, which we use to identify the research gap. Section 4 links this gap to information quality and the potentials of big data analytics. Subsequently, we detail our research methodology. In our main contribution, we establish hypotheses, perform operationalization as well as construct a structural equation model, and present the results of a survey. Finally, we analyze the findings and discuss them in light of related work.

2 Sustainability reporting and its credibility problem

2.1 Sustainability reporting

Sustainability. Sustainability is defined as “development that meets the needs of the present without compromising the ability of future generations to meet their own needs” (WCED 1987, 41). Sustainable development therefore implies balancing three dimensions: economic, ecologic, and social (Zimmermann 2016, 5). The objective of the economic dimension is to ensure long-term returns. To achieve this, the use of resources is necessary. Sustainable management constrains this use through the other two dimensions and, thus, provides a responsible scope for action. The ecological dimension emphasizes the elusive value of nature and the finiteness of nature’s resources. Its focus is on the protection of the ecosystem as the basis of human existence. The social dimension focuses on distributive justice. It is about equal opportunities and access to resources for countries, societies, and generations (Bansal 2005, 198f).

Concept of corporate social responsibility. The concept of corporate social responsibility (CSR) is the entrepreneurial answer to integrate the idea of sustainability into economic action (Zimmermann 2016, 17). It refers to a broad spectrum of activities with which companies assimilate social and sustainability interests into their business activities, mostly on a voluntary basis (ISO 2010, 3; Zimmermann 2016, 11). In doing so, any enterprise must take into account the expectations of stakeholders, that is any group or individual who is affected by the achievement of corporate objectives (Bundesministerium für Arbeit und Soziales 2011, 12; Freeman 2010, 46f). Stakeholder contribution to business activities is, therefore, essential. Business must engage with those directly and indirectly involved (Bundesministerium für Arbeit und Soziales 2011, 26f).

Sustainability report. According to the Global Reporting Initiative (GRI) (KPMG 2013, 12), a sustainability report is the key platform for communicating qualitative

and quantitative corporate sustainability performance and impact in all three aforementioned dimensions—positive as well as negative (GRI 2016, 3 and 13). According to Fifka (2014) the information contained in the reports serves internally for more sustainable control and the optimization of processes. Externally, it supports the communication with stakeholders. A sustainability report can be part of a larger integrated report also containing financial performance indicators.

In the following, we focus on sustainability reports of non-financial segments in all three dimensions. Exemplary content are explanations or key figures of identified future trends, strategic orientations and intentions, statements on corporate research and development, corporate energy consumption, occupational accidents, or social projects supported.

Credibility doubts. Frameworks for reporting such as the GRI, which have been largely compliant with the EU directive 2014/95/EU before it became mandatory legislation, are widely used. Yet, at times there are doubts as to whether the published information of the reports is true to the word (Fifka 2014, 5; Weber 2014, 97). On the communicator side, there is temptation to achieve advantages from positively communicated sustainability activities without carrying out the necessary efforts (Zimmermann 2016, 10 and 17). This deception is called *greenwashing*. By deliberately exaggerating or misreporting, companies try to be perceived in a more environmentally friendly and responsible manner (Weber 2014, 104). The intent is to gain advantages such as improved reputation, an improved awareness of the customer to purchase, or an improved motivation and recruiting of employees (Fifka 2014, 12; Weber 2014, 103). On the other hand, the credibility is at risk due to (unintentional) deception if information is left out. A recipient of information itself does not distinguish between these possibilities in his or her perception of truth (Köhnken 1990, 4).

2.2 Perceived credibility

Credibility. In scientific literature, one can find two views on credibility. According to the *communicator-centered view*, a message is credible if a communicator passes on information that he believes to be accurate. Thus, the communicator does not have any intention to mislead the recipient (Köhnken 1990, 4). Conversely in the *recipient-centered view*, a message is not considered credible based on the sender's intention. It defines credibility as result of a subconscious appraisal process by the recipient as an individual. This is applicable to persons or institutions, to spoken words, to texts, or even to visualizations (Nawratil 1999, 15).

Perception of truth. Accordingly, the perception of truth is not directly related to the truth value of a message. It is about the subjective belief that the message is true (Spelthahn et al. 2009, 62). Nevertheless, both views have in common that the perception of truth can be explained as one's (i.e., the communicator's or recipient's) subconscious comparison between the reality as an intended ideal of truth and the given set of information. Reality itself can be understood as an objective external world, which enables a subjective interpretation with a structured and systematic character by means of certain characteristics or stimuli (Früh 1994, 22ff and 54). Hence, a credible message is subconsciously always associated with a

perceived truth value, which assumes a precise representation of the assumed real-world (model).

Information asymmetries. The question of credibility itself arises when information becomes relevant for decisions or actions but is not known yet from personal knowledge or experience (Köhnken 1990, 1). In this sense, a sustainability report represents the starting point of an information asymmetry between the sender and the recipient on whether the report contains the (full) truth or not. This represents a state of uncertainty. At the same time a company aims to present information credibly as the recipient will consider it in his decision-making if he perceives it as credible (Bentele 1988, 407). Conversely, the potential benefits of a sustainability report are linked to a credible perception of its content; otherwise, its creation was for nothing or may even be harmful.

Agency dilemma. Agency theory is referred to as the scientific explanation for the state of insecurity due to information asymmetries with two parties of different intents. One party (principal) is dependent on the actions of another party (agent) (Pratt and Zeckhauser 1985, 2). Due to an information disadvantage, the principal cannot effectively control the agent and does not have direct control over him. The agent will often deviate from the principal's expectations in favor of his own interests. The greater the asymmetry, the more difficult it is to counteract the principal. The aim of the theory is an incentive-compatible remuneration system, which motivates the agent to make a decision in the sense of the principal, so that direct control is not necessary (Kleine 1995, 1f and 29ff).

In the context of this paper, the publishing company is the principal, and the respective recipient is the agent. An agent will initially be critical of the published content, in this case a sustainability report (Weber 2014, 97; Fifka 2014, 5). This raises the question of credibility. The problem has consequences for both parties. It begins with an untruthful perception of the information in a sustainability report. The recipient will not take the new knowledge into (future) decision-making process(es) (Bentele 1988, 407). Thereby, the recipient's lack of information on certain issues reduces the quality of his decision-making with respect to the company. On the company side, the sustainability report becomes worthless or even has a detrimental effect (Weber 2014, 97) towards its relation to the agent. To counteract these problems, a company publishing sustainability reports must focus on credible communication of content. Therefore, the aim must be to close the confidence gap, also known as the credibility gap, which is derived from the information asymmetries of the agency dilemma and is closely linked to credible communication.

2.3 Credibility of communication

Communication. Communication and sustainability reporting are directly linked to each other (Lock 2016, 425). To overcome the credibility gap, one needs to gain an explicit understanding of the key points responsible for the perception of a recipient on corporate communication. On the one side, the communication act itself always consists of three entities: sender (communicator), message, and recipient. On the other side, (moral) legitimacy is the central problem of corporate communication

and can be traced back to the above-mentioned information asymmetries (Bentele and Seidenglanz 2015, 411f). Legitimacy itself is defined as the perception, whether the company is in line with some socially constructed system of norms, values, beliefs, and definitions (Suchman 1995, 574).

Credibility. The *Habermas Theory of Communicative Action* discusses the sender, the recipient, the message, and the legitimacy of the company (Habermas 1984). The theory takes the practical and theoretical meaning of communicative action of the modern society into account and tries to (theoretically) solve the questions of truth, truthfulness, and normative justice by examining different meanings of rationality (Habermas 1984). For a discourse leading to a credible interpretation, the theory postulates that all participants have to communicate intelligibly, honestly, truthfully, and normatively correct to reach mutual understanding and agreement among all participants (Lock 2016, 422f). While credibility describes the believability of a source or message, rationality characterizes those who are agreeable to reason. In deliberative discourse, the best argument wins and prevails as a consensus (Habermas 1984, 96). Despite its original context of speech situations, the theory is applicable to (written) communication on sustainability issues as it has a political-normative character (Lock 2016, 423ff). Hence, the theory can be regarded as an ideal to strive for in the context of sustainability reporting (Lock 2016, 415).

Properties of credible communication. In his theory, Habermas prescribes the following four properties, which must be satisfied to perceive an interpretation of a message as credible (Habermas 1984, 329):

1. *Sincerity* Statements are reproduced honestly.
2. *Truth* Statements are in line with objective truth.
3. *Normative rightness* Statements are morally appropriate to society's requirements.
4. *Intelligibility* Statements are formulated intelligibly.

We regard *intelligibility* of the formulated statements as a foundational prerequisite. It is the only way language itself can be used as a medium to allow a rational assessment of honesty, truth, and moral appropriateness. This counteracts possible misunderstandings and perceived falsifications (Habermas 1984, 88). Since any corporation seeks its published information to be perceived as credible, we assume this dimension always to be met. Using the three further dimensions, we examine and structure knowledge about the credibility gap to extract possible starting points.

3 State-of-the-art of research on the credibility gap

3.1 Current state of research

The *credibility gap* describes a lack of confidence in the abilities and intentions of the publishing company from the viewpoint of stakeholders (Dando and Swift 2003, 196ff). It has not yet been closed (Lock and Seele 2016, 192; Knebel and Seele

2015, 196). Hence, a general statement about credibility in sustainability reports cannot be made. However, in principle it is undisputed that it poses a problem (Milne and Gray 2013, 21; Sethi et al. 2015, 61; Knebel and Seele 2015, 197). So far, the focus of research has been largely on studying the dissemination of sustainability reports, the characteristics of the publishing companies (such as size, country, industry), and the impact on financial indicators. The empirical findings of Lock and Seele (2016) indicate that sustainability reports tend to become more credible in recent years. Findings on perceived credibility from the recipient's point of view are rather rare (Lock and Seele 2016, 186).

Following Liljenström and Svedin (2005), we have categorized topics on countering the credibility gap in sustainability reporting in three core levels of a corporate environment: micro, meso, and macro. Macro or 'macroscopic' describes the entire ecological system. Meso or 'mesoscopic' describes a group or population within the macro system. Finally, micro or 'microscopic' describes the individual. Due to the importance of reporting standards and the undergoing changes of legal requirements for sustainability reports, we distinguish *external influencing factors* and *standards and legal requirements* on the macro level. On the meso level, we have subsumed all actions and decisions of corporate external services as *external audits*. On the micro level, we summarize corporate *internal potentials*, which infer the possibility of a closer cooperation or access to information for involved parties. An overview of the mentioned topics and their activities is given in Table 1. The table also includes references for proven and refuted impacts of the subtopics.

External influencing factors. As *external influences* on the perception of credibility, *culture* (Fifka 2013, 24f), *industry differences* (Fifka and Drabble 2012, 461 and 464–468), *size* (Lock and Seele 2016, 188), and *reporting experience* (Albertini 2014, 237–252) have been investigated. The studies only confirm an influence of *culture* (Fifka 2013, 24f; Freundlieb et al. 2014, 32–41).

Depending on their own cultural background, recipients expect the publication of certain sustainable activities (Adnan 2009, 9 and 14ff; Fifka 2013, 24f). Differences between sectors are also suspected. However, it was not possible to establish a link between the increased credibility of reports from environment-related areas and reports from other industries (Fifka and Drabble 2012, 466ff; Lock and Seele 2016, 189 and 192f). In contrast, the size of a company has a positive influence on reporting (Fifka 2013, 24ff). Even so, according to Lock and Seele (2016), this does not lead to increased credibility. Reporters, however, go through a learning curve that can have a positive effect on the relevance of the published information and its credibility in general (Albertini 2014, 237–252).

Standards and legal requirements. Studies on the impact of (national) *legal requirements* (Habek and Wolniak 2016, 412ff; Lock and Seele 2016, 189 and 193) and the use of (common) *reporting standards* (Knebel and Seele 2015, 199f and 204f; Lock and Seele 2016, 188 and 193) reveal that both contribute to an increased credibility. The extent of this depends on the associated review methods of a company.

Ioannou and Serafeim (2014) show that there is general improvement due to legal requirements. According to Schaltegger (1997), government regulations do not necessarily lead to an increase in quality. Vormedal and Ruud (2009) attribute this

Table 1 Categorization and summary of current findings to counter the credibility gap

| Topic | Subtopic | Impact | References |
|--|-----------------------------|--------------------------|---|
| External influencing factors (macro) | Culture | + | Adnan (2009); Fifka (2013); Freundlieb et al. (2014) |
| | Industry differences | 0 | Fifka and Drabbler (2012); Lock and Seele (2016) |
| | Size | (+) | Fifka (2013); Lock and Seele (2016) |
| | Reporting experience | (+) | Albertini (2014) |
| Standards and legal requirements (macro) | Legal requirements | + | Schaltegger (1997); Vormedal and Ruud (2009); Ioannou and Serafeim (2014); Habek and Wolniak (2016); Lock and Seele (2016) |
| | Reporting standards | + | Morhardt et al. (2002); Adams and Evans (2004); Knebel and Seele (2015); Lock and Seele (2016) |
| External audits (meso) | Reasons external review | + | Blackwell et al. (1998); Carey et al. (2000); Dando and Swift (2003); Hodge et al. (2009) |
| | Review standard | + | Manetti and Becatti (2009); Hodge et al. (2009); Frost and Martinov-Bennie (2010); Knebel and Seele (2015) |
| | Extent of the review | (+) | Manetti and Becatti (2009); Hodge et al. (2009); Frost and Martinov-Bennie (2010); Knebel and Seele (2015); Hsueh (2018) |
| | Selection of auditor | + | Wallage (2000); Ball et al. (2000); Dixon et al. (2004); Hodge et al. (2009); Simnett et al. (2009); O'Dwyer et al. (2011); Perego and Kolk (2012); Ackers and Eccles (2015); Gürtürk and Hahn (2016) |
| | Independence of auditor | (+) | Ball et al. (2000); Hodge et al. (2009); Simnett et al. (2009); O'Dwyer et al. (2011); Gürtürk and Hahn (2016) |
| | Stakeholder involvement | + | Thomson and Bebbington (2005); Perrini (2006); Manetti and Becatti (2009); Manetti (2011); O'Dwyer et al. (2011); Manetti and Toccafondi (2012) |
| | External rating | (+) | Chatterji and Levine (2006); Robinson et al. (2011); Windolph (2011) |
| | Internal potentials (micro) | Sustainability committee | + |
| | NGO cooperation | + | Amran et al. (2014) |
| | Internal audits | + | Trotman and Trotman (2015); Gürtürk and Hahn (2016) |
| | Integrated report | 0 | Adnan (2009); Lock and Seele (2016) |
| | Length of report | 0 | Fifka and Drabble (2012); Lock and Seele (2016) |
| | Balance of information | + | Guthrie and Farnetti (2008); Milne and Gray (2013); Mishra and Modi (2013); Lock and Seele (2016) |

Impact: '+' = positive given impact proven; '(+)' = positive impact under discussion; '0' = no positive impact proven

to limited political and social options to influence. Legal requirements only contribute to increased quality and credibility if the conditions and governmental controls are appropriate (Habek and Wolniak 2016, 414).

A major problem of non-financial reporting is its incompleteness (Adams and Evans 2004, 104f). Morhardt et al. (2002) indicate that the guidelines of the GRI lead to expanded reporting requirements and improved completeness. Lock and Seele (2016) verify that the use of standardized GRI guidelines leads to more credible reports than non-standardized reports. On the other hand, they refuted the assumption that a stricter compliance with the guidelines leads to reports that are more credible.

External audits. An *external audit* is positively correlated with the perceived credibility of the respective recipients for several reasons (Hodge et al. 2009, 179ff). In addition, a *review standard* increases the trustworthiness of the review process itself, whereas there is no general consensus on the *extent of the review* (Knebel and Seele 2015, 201f and 206f; Hsueh 2018, 556ff). The *selection and independence of the auditor* is also important (Hodge et al. 2009, 181–190). *Stakeholder involvement* in the review process seems to have a positive impact as well (Manetti and Becatti 2009, 292–295). Considering *external ratings* opinions differ (Chatterji and Levine 2006, 31ff; Robinson et al. 2011, 498–503).

Blackwell et al. (1998) identified the reduction of information asymmetries to lenders as the main *reason* for an external audit. Carey et al. (2000) confirm that external audits lead to better conditions and lower monitoring efforts by financial institutions. Ball et al. (2000) found that no investigated report was verified independently in its entirety. The problem is payment of the reviewer by the communicator (principal) rather than by the recipient (agent). Dando and Swift (2003) also confirm that a certified report is no guarantee for credible perception.

Review standards counteract heterogeneity and arbitrariness of the auditing process. Hodge et al. (2009) argue that the declaration of the audit statement does not lead to a more credible perception. For Frost and Martinov-Bennie (2010), this is due to a considerable lack of understanding among the recipients of the report. Related to this, the *extent of the review* seems to be related to legislation. However, the selection and calculation of many indicators in a sustainability report vary, which makes them difficult to assess (Knebel and Seele 2015, 198ff).

The *selection and independence of the auditor* has a direct effect on the quality of the reports (Perego and Kolk 2012, 176–186). According to Wallage (2000), professional auditors lead to higher quality reports and better credibility. Ball et al. (2000) also share the same view. Perego and Kolk (2012) as well as Ackers and Eccles (2015) confirm this finding. However, it is contradicted by the view of Dixon et al. (2004) and Hodge et al. (2009). For them, specialized consultants have a higher level of competence and provide a more balanced audit explanation. Simnett et al. (2009) and O'Dwyer et al. (2011) confirm this view.

For Manetti and Becatti (2009) further problems lie in the insufficient *stakeholder involvement*. Thomson and Bebbington (2005) demonstrate a link between reporting quality and stakeholder engagement. Nonetheless Perrini (2006) and Manetti (2011) prove that companies have so far been reluctant to involve stakeholders in the decision-making process on the content of sustainability reports.

This can have advantages as shown by O’Dwyer et al. (2011) as well as Manetti and Toccafondi (2012).

The improvement through *external ratings* such as ranking lists, awards, and sustainability indices is under discussion. They should serve as a neutral instance with own evaluation systems and criteria (Chatterji and Levine 2006, 31f) to be an intermediary between corporations and their stakeholders (Robinson et al. 2011, 495). Despite good intentions, there is criticism since they are depending on the disclosure of quality information by the companies themselves as the issuing bodies do not make their own measurements (Chatterji and Levine 2006, 32f). This has a weakening effect on their positive impact on credibility (Windolph 2011, 47f).

Internal potentials. Further research has been carried out on the impact of *internal potentials* of organizations. Here, positive effects of an own *sustainability committee* (Adnan 2009, 10 and 13ff; Amran et al. 2014, 222f and 226–230), a *Non-governmental Organization (NGO) cooperation* (Amran et al. 2014, 223 and 226–230), and *internal audits* (Trotman and Trotman 2015; Gürtürk and Hahn 2016) of non-financial key figures could be demonstrated. This also applies to the *balance* of published positive and negative information (Guthrie and Farneti 2008, 363ff; Mishra and Modi 2013, 434 and 441–446). Despite that, *report length* (Lock and Seele 2016, 187f and 191ff; Fifka and Drabble 2012, 461 and 465) and the type of *integrated report* (i.e., regular annual report with sustainability report) (Lock and Seele 2016, 188f) do not improve the perceived credibility. In the case of the latter, the credibility of the report even worsens (Lock and Seele 2016, 192f; Adnan 2009, 7 and 13ff).

3.2 Identified research gaps

When analyzing the current state of the art using the remaining three properties of credible communication, we found that both, the dimensions of *sincerity* and *normative rightness*, have been well covered in attempts to improve perceived credibility and to close the credibility gap. The dimension of *truth* seems to drag behind.

Sincerity. The initiation of *sustainability committees* as well as the *balance* of the content has a positive effect on sincerity. Sustainability committees ensue intensified efforts of employees involved in sustainability; balance is explained through a more honest representation of information. *External audit standards* and the *involvement of stakeholders* seem to have positive effects as well. External audits limit the scope of action of the publishing company. The involvement of stakeholders leads to an accelerated need for explanations for the intended dissemination of information.

Normative rightness. *Reporting standards*, *external audits*, and the build-up of *internal knowledge* have a positive impact on normative rightness. This can be attributed to the improved competence for the implementation of an appropriate sustainability report. In combination with the influences of *legal requirements* and the *involvement of stakeholders* into the creation and review process, the dimension seems to be well covered.

Truth. We observe a research gap in methods to improve the objective truth as well as the perception of truth as we only found few attempts dealing with this topic. Therefore, in the following we use it as the focus for our own research. As a first step, we need to examine the relationship between the perception of the objective truth by the recipient and information quality as the key factor for representing and improving objective truth as the actual state of reality in sustainability reports.

4 Bridging the credibility gap through information quality

4.1 Link between truth and information quality

A comparison between the psychological construct of credibility and that of information quality shows that it is difficult to make a clear cut. Recipients do not distinguish between credibility and quality of given information (Wirth 1999, 57f). Both constructs are multidimensional (Bentele 1988, 421) with similar to congruent criteria for their operationalization. This also applies to their evaluation: the better an individual quality assessment for presented information is, the more likely it is that the content will be used (Wolling 2004, 174; Früh 1994, 22ff and 54f).

Information quality. The quality of information is essential for the realistic and error-free reproduction of information. *Quality* is defined as an individual rating criterion in the context of quality management and assurance. The measurement is based on the ability of a product to satisfy declared or implied needs, based on its totality of characteristics (ISO 2015). *Information* is a multi-dimensional construct that can be described by means of layers. On four hierarchy levels, characters with syntax form data, data embedded in a context becomes information, and networked information come to be knowledge (Krcmar 2015, 11f). Deduced from this, high decision quality is always based on a high quality of information, which is based on high-quality data. In line with this insight, we speak of ‘fitness for use’, the suitability of information to the respective application context (Wang and Strong 1996, 6).

Information quality framework. In the following, we use the framework of Wang and Strong (1996). They have defined a set of dimensions for information quality with criteria to measure. A critical review of the framework has been mandated by the German Government and conducted by the *Deutschen Gesellschaft für Informations- und Datenqualität (DGIQ)* (Rohweder et al. 2015). Their result can be found in Table 2; an asterisk marks interpretational deviation from the original of Wang and Strong.

As introduced above, we focus our research on the truth dimension. Hence, we deem two dimensions of information quality to be of particular importance: the *intrinsic dimension* and the *contextual dimension* (Bentele and Seidenglanz 2015, 421). The dimensions of *representational* and *accessibility* are associated with the other dimensions of Habermas’ theory (in particular intelligibility). Hence, we made the conscious decision to not consider them for the remainder of our research.

Intrinsic information quality. *Sustainability committees*, *NGO cooperation*, and *internal audits* have a positive impact on intrinsic information quality. As a

Table 2 Criteria for the quality of information (Wang and Strong 1996; Rohweder et al. 2015)

| Dimension | Criteria | Explanation |
|------------------|--|--|
| Intrinsic | Information must have a high data value itself | |
| | Reputation | Reputation of a high level of trustworthiness and competence of the source, the transport medium, and the processing system through repeated positive experiences with similar information |
| | Accuracy* | Consistent with the real-world in terms of accurate, correct, reliable, and certified error-free data |
| | Objectivity | Strictly factual and impartial |
| | (Data-)Believability | Certification shows high quality standard of information processing or high effort for information acquisition and dissemination |
| Contextual | Information must be of high-quality in its context | |
| | Timeliness | Contemporary mapping of the properties of the object |
| | Completeness* | Required information scope and detail from data basis possible |
| | Appropriate amount of data | Available amount of information meets the requirements |
| | Value-added | Economically profitable for decision making |
| | Relevancy | Providing necessary and useful information for users |
| Representational | Information must be easily and comprehensibly cognitively understandable | |
| | Interpretability | Evident and purposeful for the user |
| | Ease of understanding | Required information presented concisely and comprehensibly |
| | Representational consistency | Represented continuous homogeneously |
| | Concise representation | Unmistakable for different users |
| Accessibility | Information must be easily retrievable and editable with the company's information systems | |
| | Accessibility | Retrievable by simple procedures and directly |
| | Access security | Easy to change and multilateral usable |

consequence of the resulting increase in know-how in terms of competence and trustworthiness, the criterion of reputation seems to be satisfied. This also applies to the criterion of objectivity due to the *involvement of stakeholders* with major interests of sustainability. (Data-) Believability, in the sense of a high expenditure for data acquisition and processing as well as a certification, can also be confirmed. This is based on a combination of the measures outlined above with the further, limited actions of an external audit by a *high-quality auditor* and the *involvement of stakeholders*. On the other hand, the criterion of accuracy, in sense of conformity with reality due to the precision of data, is improved only in a limited fashion. Despite increased efforts, the information quality of sustainability reports still only moderate. Thus, the intrinsic dimension seems to be satisfied except for the criterion of limited accuracy.

Contextual information quality. Previous attempts to improve contextual information quality have been less successful. Similarly, *sustainability committees*, *NGO cooperation*, and *internal audits* strengthen sustainability reports. Due to the increase of expertise in the sense of competence and suggested neutrality, the criteria of completeness and appropriate amount of data improve. However, we question the satisfaction of the criteria due to the discovered quality deficiencies. In addition, the limited impact of *external audit* methods, high-quality *auditors*, and *stakeholder involvement* can be expected only to lead to a limited increase in relevancy and timeliness. The criterion of added-value seems to be of lesser importance as it has not been considered so far. Thus, the contextual dimension does not even seem to satisfy one criterion in its entirety.

Addressing the deficits. Due to the number of deficits in both dimensions, we limit ourselves to and, thus, focus on a selection of meaningful criteria. Therefore, we have used the GRI v4 guidelines as the most commonly used framework of sustainability reporting worldwide (KPMG 2013, 12) for prioritization. Their recommendations demand stakeholder inclusiveness, sustainability context, materiality, completeness, balance, comparability, accuracy, timeliness, clarity, and reliability (GRI 2016, 9–18). Related research has shown that stakeholder inclusiveness, sustainability context, balance, comparability, and clarity of presented information can be regarded as part of the dimension of *intelligibility* and that reliability can be achieved through *external audits*. The criteria of materiality, completeness, accuracy, and timeliness remain problematic.

Three of the four criteria of GRI (GRI 2016, 9–18) are consistent with those from Wang and Strong (1996) as found in Table 2 *Materiality* is exclusive to GRI and provides recommendations for the prioritization of topics in reports (GRI 2016, 11f). At first sight, this might suggest a trade-off with the criterion of *completeness*. Yet, it is not, as completeness has to be understood as scope, boundary, and time of the report incorporating the measure of prioritization as well as the practices of information collection (GRI 2016, 12f and 17). Therefore, it is a prioritization principle that is applied when compiling the measures to include in the report. It does not affect the information quality of the reported measures once they have been selected. Hence, we refrained from including materiality as a criterion.

The importance of the three remaining criteria of *timeliness*, *completeness*, and *accuracy* is supported by a survey of Michnik and Lo (2009). The author examined the relevance of the above-mentioned four-dimensional representation of information quality and their criteria for data users. Therefore, we adopt *timeliness*, *completeness*, and *accuracy* as criteria for the representation of information quality.

4.2 Potential of improvement with big data analytics

The basic assumption is that credible perception correlates with the quality of information. As argued, this in turn is dependent on the quality of data. It allows us to conclude that there is potential for improvement through effective and efficient data processing. Due to the emergence of data identified as of big data, there is now a major trade-off between size, time, quality, and cost of information generation that cannot be dealt with in terms of traditional business intelligence capabilities (Schön

2016, 19f) and may even lead to a situation where companies are confronted with a data deluge (Müller et al. 2018, 489). Thus, we assume that the perception of information in sustainability reports can be improved through big data analytics.

Characteristics of big data. Ylijoki and Porras (2016) provide an extensive, up-to-date survey of big data definitions. They identified 17 interpretations from a total of 479 scientific articles. The result confirms the importance of three characteristics: *Volume* (95%), *Variety* (89%), and *Velocity* (74%). *Volume* denotes an unusually large amount of data (Géczy 2014, 98; Tole 2013, 32). *Variety* covers the diversity of data sources and formats and in particular enables the processing of unstructured data (Géczy 2014, 98f; Tole 2013, 33). *Velocity* refers to the speed of data modification and evaluation (Géczy 2014, 98; Tole 2013, 32). In the following, we use the term *data work* to describe different types of processing in big data analytics (e.g., in scope and frequency) leading to sustainability reports.

Van Altena et al. (2016) come to the same conclusion. Deviations can be found in the dimensions *Veracity* (23%) and *Value* (27%). These appear predominantly in newer interpretations and should, therefore, be taken into account. *Veracity* includes the reliability of data through extensive testing routines (Schön 2016, 304; Van Altena et al. 2016, 9). *Value* refers to the value of the use of technology from an economic point of view (Tole 2013, 32f; Van Altena et al. 2016, 9).

Reviewing the characteristics of big data, the size of the data set is not the defining criterion for big data. However, at least one of the three main characteristics (i.e., volume, variety, velocity) should be linked to its economic use.

Big data analytics. Big data require “the use of powerful computational techniques to unveil trends and patterns within and between these extremely large socioeconomic datasets” (George et al. 2014, 321). We term these techniques *big data analytics*. Big data analytics provides a physical (hardware technology) and digital (software technology) materiality representing stable properties across contexts and time. Examples for physical technologies include in-memory databases (chiefly volume, velocity) or in general contemporary compute, storage, and network capabilities. Examples for software are more diverse and include NoSQL databases such as Apache Cassandra or Amazon Dynamo (volume), event stream processing engines such as Esper (velocity), or statistical software such as R (variety). They provide affordances as potentials for action to process data and create comprehensive information for media such as sustainability reports (Lehrer et al. 2018, 428f). A further overview of exemplary physical and digital technologies can be found in Lehrer et al. (2018).

Traditional business intelligence software and hardware technology focuses on providing management with a comprehensive set of key performance indicators based on consistent data to assess past and current business performance to improve efficiency of existing operations. Analytics software and technology (such as big data analytics) further enables analytics-savvy managers and data scientists to explore, discover, and predict new ways to conduct business (Abbasi et al. 2016, viif). The exact extent of processing for example high-volume, high-velocity, or high-variety data remains unspecifiable and depends on the respective context.

Big data analytics, for instance, can give real-time access to analytics of trace data using sensor networks, scalable in-memory access to large amounts of data

points, or text and sentiment analysis of heterogeneous external reports or online conversations such as opinions or feedback. Considering all of these applications, big data analytics offer diverse opportunities for innovation and business transformation. To capitalize on these affordances, one needs to use big data analytics appropriate to one's context.

Summarizing, the core aim of big data analytics is to improve insight, decision-making, and process automation from the analysis of (complex) data sets under economically feasible conditions. Technological choices must be made according to one's context and the extraction of the intended data value must be carried out while ensuring high data quality. In terms of sustainability reports, the result is an increase in information quality through extensive data work in any or multiple of the mentioned dimensions, which leads to an improved truth value of the content. This in turn should improve credibility.

5 Research methodology

Due to the lack of scientific knowledge and novelty of the topic, we chose the survey of experts with IT- and data-affine persons as the method of investigation (Przyborski and Wohlrab-Sahr 2014, 124ff). For its development, we carried out a quantitative and qualitative survey and evaluation. We used the quantitative survey to pinpoint to what extent there is an agreement on the perception of the credibility of recipients of sustainability reports with regard to information quality criteria. The qualitative survey supported the item selection of the survey.

Statistical evaluation. We chose the causal analysis for the statistical evaluation. It allows the investigation of collected data sets for suspected cause–effect relationships. Mathematically, this is based on a combination of three statistical approaches: factor and path analysis and multilinear regression (Kühnle and Dingelstedt 2014, 1017–1028). We used the variance-analytical approach for the estimation of the structural equation model with its multi-variable system of equations. The model is highly prognosis-oriented, with the aim of explaining latent and/ or associated indicator variables. Both reflective and formative measurement models can be used for the model structure (Chin and Newsted 1999, 314). This allows us to prove the assumed interrelationships and assess potential improvements through big data analytics.

We selected the partial least squares (PLS) as the statistical evaluation method. This method is suitable due to its focus on predictions, lack of well-founded measurement and construct theories, and a lack of covariance-based independence of the observed values (Chin and Newsted 1999, 314). The use of PLS requires seven steps (Weiber and Mühlhaus 2014, 325): hypothesis and modeling, construct conceptualization, construct operationalization, evaluation of the measurement models, model estimation, evaluation of the overall model, and result interpretation.

Progression of research. First, we derived hypotheses and justified them based on the information quality criteria presented above. We processed cause-effect relationships with the latent and manifest variables in a structural equation model. Subsequently, we determined suitable indicators, which can be observed and

evaluated, for all not directly observable latent variables and, thus, for all theoretical constructs to represent them in the best possible way. The operationalization of the latent variables was done by recording and measuring the hypothetical constructs based on indicators and measurement rules. For this purpose, we generated a rough classification of potential measurement indicators, defined the measurement concept, and designed the measurement specifications. We then tested the measurement models for their quality based on several reliability and validity tests. For the model estimation, we first cleaned the empirical data obtained during the main investigation and then applied it to the structural equation model using SmartPLS. Reliability and validity of the overall model were again tested to ensure a sufficient model fit. Finally, we examined the a priori hypotheses on basis of the empirically collected data and interpreted them.

6 Perceived credibility of corporate published information in sustainability reports

6.1 Hypothesis determination and derived structural equation model

In the following, we examine whether it is possible to improve the perception of the objective truth of published sustainability report information by the means of big data analytics. As indicated by our own analysis so far, the improved representation of reality should lead to an improved objective truth and, thus, an improved perceived credibility by the recipient. Nevertheless, the question remains to what extent the limited rationality of human action in perceiving sustainability reports supports or refutes this assumption. For the examination, we developed the following hypotheses for the selected criteria of *timeliness*, *completeness*, and *accuracy*. The items are based on the above findings from research on information quality as well as mathematical terms for their calculation from the field of data quality (for more detail see the Appendix). All of our hypotheses aim to improve the perception of objective truth by the recipient as (a) we cannot observe and measure objective truth itself and (b) a true report, which the recipient does not consider credible, is futile.

Timeliness. The common ground of measurement methods for determining the validity of data can be found in the application of probability theory (Heinrich and Klier 2015; Hinrichs 2002). The value of data decreases exponentially over time (Heinrich and Klier 2015, 91f). The timeliness of data remains dependent on the time of delivery to the respective recipient. Consequently, this can only be validly determined by the recipient (Wang and Strong 1996, 7). This leads to the following hypothesis H1: ‘The better the expected timeliness of data, the better the perceived objective truth.’

Velocity. Due to changes in the environment (Seufert 2016, 40) an increase of data value volatility can be observed. This means that there is a reduced half-life of data. This increases the technical speed requirements for the underlying data work (Vargas-Solar et al. 2016, 2–12; Seufert 2016, 41 and 48–54). A high up-to-date value implies new hardware and software technologies as promised by big data

analytics. Velocity is necessary per a certain level on the timeliness of data resulting in hypothesis H2: ‘The greater the need for the big data characteristic velocity, the better the expected timeliness of the data’

Completeness. A value for the completeness of data compared to reality can be determined at the level of the database (Aljumaili et al. 2016; Heinrich and Klier 2015). Incomplete entries represent unknown or missing data (Batini et al. 2009, 7). Completeness represents an expected extent to which relevant data for a specific scope of application are available for big data analytics (Aljumaili et al. 2016, 244). Conversely, we can assume as well that the more comprehensive the analysis and assignment of data values, the better its completeness. Again, the quality can only be determined by the recipient (Wang and Strong 1996, 7). Consequently, hypothesis H3 is defined as follows: ‘The better the expected completeness, the better the perceived objective truth.’

Volume. Nowadays, large amounts of data have to be evaluated (Seufert 2016, 40f). This cannot be realized in an economic fashion with traditional IT standards. Physical and digital big data analytics technologies based on volume promise a remedy. Based on this, we formulate hypothesis H4: ‘The greater the need for the big data characteristic volume, the better the expected completeness.’

Variety. In addition to the amount of data, the origin and structure of data change. Nowadays, more and more video and audio material, browser data, simulations, or gyrosopic data have to be evaluated (Vargas-Solar et al. 2016, 2–12). In the past, IT methods have not been implemented for this purpose and reach physical and economic limits. Problems arise primarily with regard to the evaluation of semi- and unstructured data, which necessitates complex pre-treatments (Schön 2016, 298). More and more information is becoming available, which calls for innovative technologies (Seufert 2016, 48ff). Physical and digital big data analytics technologies catering for the characteristic of variety promise a solution to this (Seufert 2016, 53f). We derive the following hypothesis H5: ‘The greater the necessity of the big data characteristic variety, the better the expected completeness.’

Accuracy. To determine the value of accuracy, one can compare data from an information system with a data entity x assumed to be free of errors to the modelled reality (Aljumaili et al. 2016; Heinrich and Klier 2015; Hinrichs 2002). Accuracy data denotes sufficient detail and exactness in the measurement and retrieval of data as well as validation routines (Aljumaili et al. 2016, 243f). However, this is also depending on the individual decision, as a (subconscious) validation process takes place due to information asymmetries (Shankaranarayan et al. 2003, 9). Therefore, the recipient must determine the value. This leads to hypothesis H6: ‘The better the expected accuracy, the better the perceived objective truth.’

Veracity. Due to a multitude of new data sources and data creators, the origin of data becomes increasingly questionable (Lukoianova and Rubin 2014, 4f). It is also important to avoid the challenges of internal and external manipulation attempts (Kepner et al. 2014, 1). This makes the process of matching reality enormously difficult. Therefore, the necessary effort for technical reconciliation and validation is growing (Schön 2016, 304). It can only be done economically justifiable on the basis of random checks (Hinrichs 2002, 87). Again, new hardware and software technologies for big data analytics promise improvements, which are reflected in the

characteristic of veracity. We derive hypothesis H7 accordingly: ‘The greater the need for the big data characteristic veracity, the better the expected accuracy.’

The hypothesis results in the following structural equation model:

Equation model. The dependencies of the perception of the *objective truth* on the information quality criteria of *timeliness*, *completeness*, and *accuracy* are our starting points (H1, H3, H6). We can also assume that they can be improved by improvement of information quality through big data analytics (represented here through its characteristics) (H2, H4, H5, H7). According to these assumptions, all links shown in Fig. 1 have a positive effect. For clarity reasons, the residual variables and the ‘placeholder variables’ for the path strengths were not included in the figure.

6.2 Construct conceptualization

For the subsequent operationalization, we first describe the latent variables and define them as precise as possible (Weiber and Mühlhaus 2014, 95–102). A summary overview of the latent variables of the structural equation model can be found in Table 3.

Construct decision levels. For the three central levels according to Rossiter (2002), the following decisions were made in our own investigation:

1. *Subject level* (target persons): target persons for the survey are employees up to 60 years of age and/or knowledge of sustainability reports. We justify this by the fact that many respondents will become potential recipients and possible

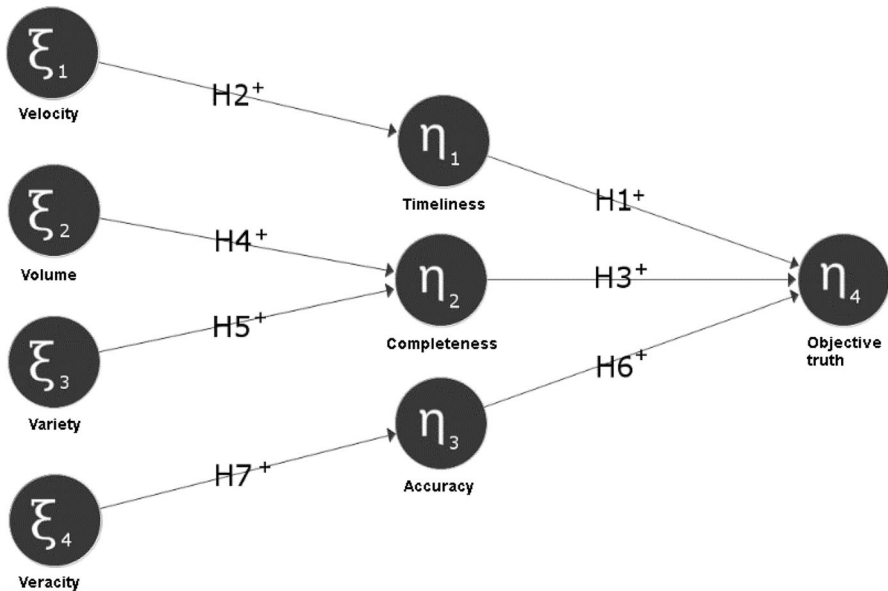


Fig. 1 Derived structural equation model without measurement criteria

Table 3 Brief descriptions of the latent variables

| Latent variable | Explanation |
|-----------------|--|
| Objective truth | Assumed current state of reality |
| Timeliness | Contemporary mapping of the properties of the object |
| Completeness | Coverage of information reach and detail |
| Accuracy | High value of data, comparison of correct reality representation |
| Velocity | Speed of data processing within data work |
| Volume | Unusually large amounts of data |
| Variety | Diversity of (new) data sources and formats |
| Veracity | Tested reliability of the (comprehensive) data |

decision-makers in the next few years. We have also selected test persons with a uniform cultural background due to the findings in Sect. 3.1.

2. *Objective level* (carrier of the assessment): participants assess the perceived information quality and perceived objective truth of different forms of big data analytics data work based on extracts from sustainability reports.
3. *Attribute level* (object properties of the appraisal): participants assess the changed conditions of the selected information quality criteria in conjunction with the perceived objective truth-value.

6.3 Construct operationalization

A data quality framework from the health sector (Canadian Institute for Health Information 2009) and topic-related mathematical calculation models (see Appendix) were used to pre-operationalize the theoretical constructs for the selected criteria of *timeliness*, *completeness*, and *accuracy*. Cf. Table 4 for the derived measurement indicators and their explanation.

Expert survey. To validate the identified measurement indicators, a qualitative expert survey was conducted with a total of 16 participants. After explaining topic and intention, a questionnaire was sent to the experts to validate the suitability of the initial measurement indicators, to determine their weighting for the respective theoretical construct, and to make proposals for further indicators. The results are also included in Table 4, where ‘ \emptyset -weighting’ is the mathematically calculated average score when ‘ \emptyset -approval’ is $\geq 75\%$. This ensures that a majority of experts support the need and legitimacy of the respective measurement indicators.

Expert discussion. Only two experts made further proposals to supplement missing measurement indicators. The proposals themselves (source reliability, non-redundancy, consistency) were thoroughly reviewed, but had to be rejected due to duplication, as they are already covered by other indicators (sources considered, validation level).

Big data analytics value determination. For the characteristics of big data analytics, no suitable studies could be found to derive measurement indicators.

Table 4 Derived measurement indicators of the information quality criteria

| Dim. | Measurement indicator | Explanation | Ø app. (%) | Ø wei. (%) |
|--------------|------------------------------------|---|------------|------------|
| Timeliness | Update frequency (U_frequen) | Time interval between measurements | 94 | 34 |
| | Aging rate | Data validity period as defined by data experts and users | 63 | – |
| | Data age (D_age) | Time between data collection and information delivery | 100 | 44 |
| | IT-speediness (IT_speed) | Time required for data processing | 75 | 22 |
| Completeness | Amount of data (S_volume) | Total number of included data records | 100 | 37 |
| | Sources considered (S_consider) | Data sources for dataset generation | 100 | 36 |
| | Formats considered | Data formats for dataset generation | 25 | – |
| | Attributes considered (A_consider) | % coverage of most important attribute values | 81 | 27 |
| | Coverage ratio | % coverage of relevant real world scenario | 63 | – |
| Accuracy | Precision (Precision) | Precision of measures and stored data values | 100 | 39 |
| | Validation level (V_level) | Scope of data and source validation | 94 | 23 |
| | Validation frequency (V_frequen) | Frequency of validation methods | 81 | 17 |
| | Error ratio (E_ratio) | Ratio of incorrect/non-existent data attributes to error-free data attributes | 88 | 22 |
| | Importance measure | % validity rate of important attribute values | 63 | – |
| | Deviation from reality | Missing/NULL values compared to reality | 63 | – |

Therefore, we did not carry out a separate expert survey. The value determination for the individual big data analytics characteristics within the survey are based on previous academic findings (Schön 2016; Van Altna et al. 2016; Vargas-Solar et al. 2016; Géczy 2014) and an expert survey according to Seufert (2016). Furthermore, due to the lack of a complete reflectivity of the theoretical constructs, the measurement concept was uniformly defined to be formative.

Operationalized equation model. As the result of operationalization, the structural equation model is extended to include measurement indicators and their weighting for the underlying latent variables. The result is shown in Fig. 2.

For the big data characteristics “PreD_V1” to “PreD_V4” a predefined value is specified. The weighting is, therefore, “1.00” each. The *objective truth* also receives “1.00”, since it contains only one measuring indicator. The respective weighting of the other variables is based on the expert survey. The designations can be found in Table 4 as “Ø wei(ghting)”.

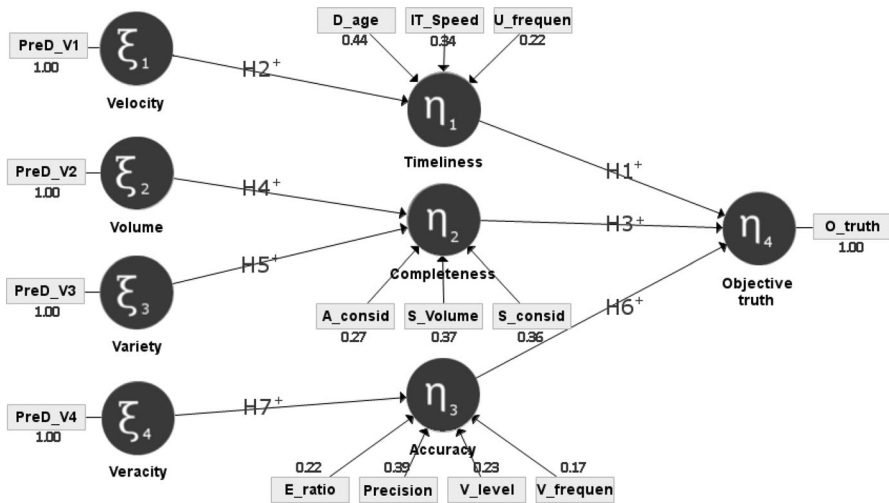


Fig. 2 Derived structural equation model with measurement criteria

The residual variables and placeholder variables for the path strengths have been omitted again for clarity reasons. Due to mathematic rounding, summations may result in values $\neq 1$.

6.4 Evaluation of the measurement models

Ultimately, we conducted a quantitative survey to examine the extracted structure equation model. Therefore, a six-tiered (Trommsdorff 1975, 93ff) Likert-scale was used as the measuring method (Weiber and Mühlhaus 2014, 43). Thus, there is a forced decision of the respondents on their tendency, which prevents undetectable reasons for the choice of averages (Weiber and Mühlhaus 2014, 42). For the grading, terms “very good” to “very bad” were used.

Survey setup. The survey itself was carried out using two questionnaires of the same type (group A and B) with different extracts from sustainability reports. A uniform language was chosen to counteract the falsification due to cultural influence (see Sect. 3.1). The questionnaire was, therefore, conducted in German language. As participants, employees were acquired from six companies ($A = 3$; $B = 3$). The sample n before adjustment was 68 ($A = 37$; $B = 31$). Exclusions of age 60+, lack of knowledge about sustainability reports, and detected anomalies in the response behavior resulted in a final sample n of 44 ($A = 26$; $B = 18$).

Survey content. As the basis for our questions, we used excerpts from actual sustainability reports by Volkswagen AG and Henkel AG and Co. KGaA (see the Appendix for details). We used these excerpts to ask questions about our variables in all three dimensions. For each dimension we formulated a text paragraph with a scenario description and asked about our variables on the aforementioned Likert scale. Once all three dimensions had been answered, we closed with a summative question on the perceived truth value. In total, we asked each question thrice

Table 5 Results from the forecast validity of the measurement indicators

| Latent variable | Measurement indicator | Forecast value |
|-----------------|-----------------------|----------------|
| Velocity | PreD_V1 | 1.000 |
| Timeliness | D_age | 0.878 |
| | IT_speed | 0.931 |
| | U_frequen | 0.701 |
| Volume | PreD_V2 | 1.000 |
| Variety | PreD_V3 | 1.000 |
| Completeness | A_consider | 0.862 |
| | S_volume | 0.900 |
| | S_consider | 0.583 |
| Veracity | PreD_V4 | 1.000 |
| Accuracy | E_ratio | 0.600 |
| | Precision | 0.844 |
| | V_level | 0.855 |
| | V_frequen | 0.596 |
| Objective truth | o_truth | 1.000 |

varying the data work involved to generate the report. For example, concerning timeliness of data we use the following three scenarios:

1. “We update our data weekly, data are evaluated every month and takes two days to process.”
2. “We update our data monthly, data are evaluated every 3 months, and take 6 weeks to process.”
3. “We update our data quarterly, data are evaluated every 6 months and take 12 weeks to process.”

Then, we asked for the recipient’s perception on the adequateness of the measurement indicators for update frequency, data age, and IT speediness.

We have varied the sequence of the data work scenario questions in the survey. Concerning completeness and accuracy, we created similar scenarios using the other indicators of Table 4: “we cover x as data sources with an inclusion of $x\%$ of possible attributes, and a total amount of x rows of data” and “we validate aspect x with a precision of x decimal places every x time units, our validation has a maximum error rate of $x\%$ ”.

The respondents evaluated their perception with the varying data work. In the same way, we examined an overall perception of *objective truth* throughout the questionnaire. Reliability and validation measures were used to ensure the consistency of the measuring instrument (Weiber and Mühlhaus 2014, 169f).

Survey validation. Forecast validity was used in the examination at indicator level. The permissible value range is a regression coefficient of ≥ 0.5 (Diamantopoulos and Riefler 2008, 1189). The calculated forecast values are available in Table 5.

Table 6 Results of the testing measures for the formative measurement models

| Latent variable | AVE | Fornell–Larcker criterion | | | | | | | |
|-----------------|-------|---------------------------|-------|-------|-------|-------|-------|-------|-------|
| | | Tim | Velo | Com | Volu | Vari | Acc | Vera | OT |
| Tim | 0.698 | 0.836 | | | | | | | |
| Velo | 1.000 | 0.621 | 1.000 | | | | | | |
| Com | 0.631 | 0.603 | 0.715 | 0.794 | | | | | |
| Volu | 1.000 | 0.621 | 1.000 | 0.715 | 1.000 | | | | |
| Vari | 1.000 | 0.562 | 0.961 | 0.710 | 1.000 | 1.000 | | | |
| Acc | 0.540 | 0.594 | 0.668 | 0.650 | 0.668 | 0.640 | 0.735 | | |
| Vera | 1.000 | 0.621 | 1.000 | 0.715 | 0.961 | 0.961 | 0.668 | 1.000 | |
| oT | 1.000 | 0.658 | 0.779 | 0.760 | 0.779 | 0.739 | 0.801 | 0.779 | 1.000 |

Values are mathematically rounded

Tim timeliness, *Velo* velocity, *Com* completeness, *Volu* volume, *Vari* variety, *Acc* accuracy, *Vera* veracity, *oT* objective truth, *AVE* average variance extracted

The result shows a high quality of the individual criteria. There are anomalies only with two measurement indicators: sources considered ($S_{\text{consid}} = 0.583$) and validation frequency ($V_{\text{frecuen}} = 0.596$). However, these remain within the defined approval range and have been confirmed by the expert survey.

At the construct level, the validity of convergence, discrimination, and nomological validity were examined. $AVE \geq 0.5$ (Fornell and Larcker 1981, 46) indicates the threshold value for good reliability in convergence validity. Values that have an average variance ratio with the latent construct that is lower than the latent construct with its own indicators ($AVE(\xi_i) > \Phi^2_{ij} \forall i, j$) (Fornell and Larcker 1981, 46) are considered to be permissible values for the discriminant validation.

As shown in Table 6, all constructs have a permissible dispersion when computing the respective AVE. Therefore, we consider convergence validity to be assured. The value of accuracy is relatively low, but nevertheless permissible ($Acc = 0.540$). One possible explanation is the high number of four measurement indicators.

In the case of discriminant validity, in addition to the fixed values for the characteristics of the big data properties allocated ex-ante, there is also an abnormality in the accuracy ($Acc = 0.735$). This shows a critical overlap with the construct of *objective truth*. One possible reason could be the sequence of the survey. Possibly, respondents are influenced by the evaluation of the measurement indicators of the preceding accuracy when grading the objective truth. The construct of objective truth is again delimitable from the other constructs. Except for the low value on the criterion of accuracy, discrimination seems to be valid.

6.5 Model estimation and evaluation of the overall model

The estimation of the structural equation model was done with the help of the adjusted empirical data from the quantitative survey. We performed an intermediate

step using the bootstrapping method (Weiber and Mühlhaus 2014, 173–198 and 323–342) due to the missing numerical values, and thus, distributions of the latent variables.

Survey reliability. The coefficient of determination (R^2) was used for checking at indicator level. The range of values is declared with $[0;1]$ and should be ≥ 0.19 in accordance with Chin (1998). The thresholds are confirmed by Hansmann and Ringle (2005), who make a further subdivision. They declare the range $0.19 \leq R^2 < 0.33$ as weak, $0.33 \leq R^2 < 0.67$ as medium, and $R^2 \geq 0.67$ as substantial. Considering the sources, a value from $R^2 \geq 0.33$ should be considered acceptable for our work.

As shown in Fig. 3, all latent variables are adequately covered by the measurement indicators assigned to them. Each of the values is ≥ 0.33 . This confirms it as a reliable structural equation model.

Survey results. Subsequently, the nomological validity was assessed at the construction level. For the proof of respective hypothesis, its path coefficient has to differ significantly from 0 (≥ 0.3) and must correspond to the a priori assumed direction of action (Weiber and Mühlhaus 2014, 265). Results are shown in Table 7 and correspond to the values in Fig. 3.

There is no assumed increase in the perceived objective truth through an increase in data timeliness ($H1 = 0.163$). Despite reaffirmation of an improvement in timeliness by velocity ($H2 = 0.621$) no increase in perceived objective truth, and thus credibility, can be achieved through reporting speed for sustainability reports. For completeness ($H3 = 0.354$) there seems to be a (low) increase. Therefore, an improvement of the objective truth by the connected criterion volume (= 0.430) can be assumed. However, this does not apply to variety (= 0.297). On the other hand, accuracy has a positive effect on objective truth (= 0.474). The criterion veracity

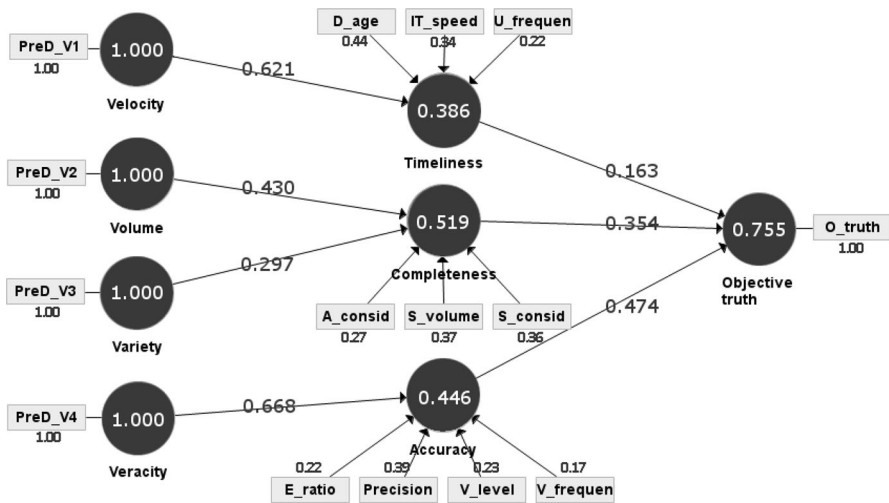


Fig. 3 Calculated structural equation model

Table 7 Result of the hypothesis test by nomological validity

| No. | Predication in short | From | After | Value | Yes? |
|-----|---|------|------------|-------|------|
| H1 | Expected timeliness of the data on perceived objective truth | Tim | σT | 0.163 | – |
| H2 | Necessity of the big data property velocity to expected timeliness | Velo | Tim | 0.621 | (x) |
| H3 | Expected completeness of the data on perceived objective truth | Com | σT | 0.354 | x |
| H4 | Necessity of the big data property Volume to expected completeness | Volu | Com | 0.430 | x |
| H5 | Necessity of the big data property Variety to expected completeness | Vari | Com | 0.297 | – |
| H6 | Expected accuracy to perceived objective truth | Acc | σT | 0.474 | x |
| H7 | Necessity of the big data property Veracity to expected accuracy | Vera | Acc | 0.668 | x |

leads to an increase in accuracy (= 0.668) and, therefore, to an increase of the credibility as well.

7 Discussion of results, limitations, and further research

The perceived credibility of sustainability reports has been a subject in research for years known as the *credibility gap*. We have devised an approach incorporating the conflict between the (ir)rationality of decisions in business and the limited rationality of human action to optimize the perception of sustainability reports, in particular in the perception of the dimension of objective truth. Despite the obvious connections between the psychological construct of credibility and the quality of the information provided, the direct impact of big data analytics remained questionable due to the limited rationality of the recipient to appreciate higher quality data. Our results indicate that the perception of sustainability reports can be improved generally with big data analytics. In doing so, we incorporate both, the communicator-centered view stating that credible communication is improved by passing on superior information and the recipient-centered view stating that credible communication is based on the recipient appraisal of the information.

Discussion. As big data analytics covers a diverse range of physical and digital technologies, their use has to be optimized for the application context. In the following, we discuss our findings on how to improve sustainability reports using big data analytics. We have found that the *timeliness* of data seems to have only a very small effect on the perception of truth by the recipients of sustainability reports. There does not seem to be any improvement in sustainability reports from big data analytics technologies such as in-memory or event processing in the *velocity* area. We assume that this is due to the fact that sustainability reports are typically periodical publications and the implementation as an updated real-time report does not provide substantial added value. Consequently, the value of real-time data for sustainability reports should always be very critically reflected and

general reports are unlikely to benefit. Nevertheless, there may be niche applications where real-time data may be valued more. In contrast, the criterion of *completeness* has a measurable impact on perception. Big data analytics with the characteristic of *volume* such as physical storage technologies or NoSQL databases improve this criterion. Recipients seem to value comprehensive data sets and analysis. *Variety*, on the other hand, as expressed through innovative data mining tools does not seem to result in a significant increase for sustainability reports. As discussed below, we assume this to be due to the area of application. Sustainability reports are based on defined measurements and performance indicators of limited variety. Recipients apparently do not value an increase of those. The criterion of *accuracy* shows the highest influence on the perception of credibility. This can be further increased by big data analytics using comprehensive and reproducible algorithms to improve the *veracity* of big data.

Our results do not confirm the assumptions of Natarajan et al. (2017) who expects that *variety* is the most significant factor of big data analytics. Natarajan et al. (2017) have made these assertions in the context of medical information quality. It is conceivable, that for medical decisions a variety of information, which points to the same diagnosis, is more important than volume or accuracy of individual data points due to the issue of differential diagnoses. Consequently, and as indicated earlier, it is important to clearly define the context in which big data analytics is employed to ensure maximal value in the delivery of information.

Limitations. There are limitations to our process of investigation during the quantitative questioning and to its scope. During the determination of the values per survey cycle, we provided explicit information on the data work undertaken for each section of the sustainability report we presented. In practice, this information typically is not available or only available to a very limited extent. Validations with other target groups also seem to be advisable to ensure that the information provided is of generalizable nature.

The Habermas Theory of Communicative Action assumes idealized conditions of discourse (i.e., an ideal speech situation) which is then immunized against repression and inequality in a special way. We have used this theory to analyze the impact of big data analytics on information quality and, thus, on the dimension of objective truth. Using the instrument of a survey, we have not analyzed the impact of further influencing factors that are inherent to our imperfect world. Hence, it is conceivable that other factors impact the recipient's perception of sustainability reports such as current (negative) media reports involving the respective companies or internet trolling.

Conclusion. We conclude that an improvement of the perceived credibility of sustainability reports is generally possible with help of big data analytics. As a recommendation to creators of sustainability reports, a focus should be placed on the information quality criteria of *completeness* and *accuracy*. Similarly, further improvement measures through future physical and digital technologies of big data analytics seem to be possible primarily with technologies that focus on the characteristics of *volume* and *veracity*, as this seems to promote the truth perception of published information in sustainability reports.

To further close the credibility gap, we suggest addressing all gaps identified by the objective truth in Habermas theory as an ideal–typical implementation of sustainability report content. Further investigations should focus on the criteria of *appropriate amount of data*, *value-added*, and *relevancy* of the contextual dimension of information quality.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Abbasi, Ahmed, Suprateek Sarker, and Roger Chiang. 2016. Big data research in information systems: toward an inclusive research agenda. *Journal of the Association for Information Systems* 17 (2): 1–32.
- Ackers, B., and N. Eccles. 2015. Mandatory corporate social responsibility assurance practices: the case of King III in South Africa. *Accounting, Auditing & Accountability Journal* 28 (4): 515–550.
- Adams, C., and R. Evans. 2004. Accountability, completeness, credibility and the audit expectations gap. *Journal of Corporate Citizenship* 14: 97–115.
- Adnan, S. 2009. Do culture and governance structure influence CSR reporting quality: evidence from China, India, Malaysia and the United Kingdom. Doctoral dissertation, University of Auckland.
- Albertini, E. 2014. A descriptive analysis of environmental disclosure: a longitudinal study of French companies. *Journal of Business Ethics* 121 (2): 233–254.
- Aljumaili, M., R. Karim, and P. Tretten. 2016. Metadata-based data quality assessment. *VINE Journal of Information and Knowledge Management Systems* 46 (2): 232–250.
- Amran, A., S. Lee, and S. Devi. 2014. The influence of governance structure and strategic corporate social responsibility toward sustainability reporting quality. *Business Strategy and the Environment* 23: 217–235.
- Ball, A., D. Owen, and R. Gray. 2000. External transparency or internal capture? The role of third-party statements in adding value to corporate environmental reports. *Business Strategy and the Environment* 9 (1): 1–23.
- Bundesministerium für Arbeit und Soziales. 2011. Die DIN ISO 26000, Leitfaden zur gesellschaftlichen Verantwortung von Organisationen“ Ein Überblick. https://www.bmas.de/SharedDocs/Downloads/DE/PDFPublikationen/a395-csr-din-26000.pdf;jsessionid=D03CEDF90C7E8A8D4DA0328277788673?__blob=publicationFile&v=2. Accessed 4 Mar 2017.
- Bansal, P. 2005. Evolving sustainability: a longitudinal study of corporate sustainable development. *Strategic Management Journal* 26: 197–218.
- Batini, C., C. Cappelletto, C. Francalanci, and A. Maurino. 2009. Methodologies for data quality assessment and improvement. *ACM Computing Surveys* 41 (3): 1–52.
- Bentele, B., and R. Seidenglanz. 2015. Vertrauen und Glaubwürdigkeit. In *Handbuch der Public Relations—Wissenschaftliche Grundlagen und berufliches Handeln*, eds. Mit Lexikon, G. Bentele, R. Fröhlich, and P. Szyszka, 411–430. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bentele, G. 1988. Der Faktor Glaubwürdigkeit Forschungsergebnisse und Fragen für die Sozialisationssperspektive. *Publizistik* 33 (2): 406–426.
- Blackwell, D., T. Noland, and D. Winters. 1998. The value of auditor assurance: evidence from loan pricing. *Journal of Accounting Research* 36: 57–70.
- Canadian Institute for Health Information. 2009. The CIHI data quality framework. https://www.cihi.ca/en/data_quality_framework_2009_en.pdf. Accessed 4 Mar 2017.
- Carey, P., R. Simnett, and G. Tanewski. 2000. Voluntary demand for internal and external auditing by family businesses. *Auditing: A Journal of Theory and Practice* 19: 37–51.
- Chatterji, A., and D. Levine. 2006. Breaking down the wall of codes: evaluation non-financial performance measurement. *California Management Review* 48 (2): 29–51.
- Chin, W. 1998. Issues and opinion on structural equation modeling. *Management Information Systems Quarterly* 22: 7–16.

- Chin, W., and P. Newsted. 1999. Structural equation modeling analysis with small samples using partial least squares. *Statistical Strategies for Small Sample Research* 1 (1): 307–341.
- Dando, N., and T. Swift. 2003. Transparency and assurance: minding the credibility gap. *Journal of Business Ethics* 44 (2): 195–200.
- Diamantopoulos, A., and P. Riefler. 2008. Formative Indikatoren: Einige Anmerkungen zu ihrer Art, Validität und Multikollinearität. *Zeitschrift für Betriebswirtschaft* 78 (11): 1183–1196.
- Dixon, R., G. Mousa, and A. Woodhead. 2004. The necessary characteristics of environmental auditors: a review of the contribution of the financial auditing profession. *Accounting Forum* 28: 119–138.
- Fifka, M. 2013. Corporate responsibility reporting and its determinants in comparative perspective—a review of the empirical literature and a meta-analysis. *Business Strategy and the Environment* 22 (1): 1–35.
- Fifka, M. 2014. Einführung—Nachhaltigkeitsberichtserstattung: Eingrenzung eines heterogenen Phänomens. In *CSR und Reporting—Nachhaltigkeits- und CSR-Berichtserstattung verstehen und erfolgreich umsetzen*, ed. M. Fifka, 1–20. Berlin: Gabler Verlag.
- Fifka, M., and M. Drabble. 2012. Focus and standardization of sustainability reporting a comparative study of the United Kingdom and Finland. *Business Strategy and the Environment* 21 (7): 455–474.
- Fornell, C., and D. Larcker. 1981. Evaluation structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* 18: 39–50.
- Freeman, E. 2010. *Strategic management—a stakeholder approach*. Cambridge: Cambridge University Press.
- Freundlieb, M., M. Gräuler, and F. Teuteberg. 2014. Corporate social responsibility reporting: a transnational analysis of online corporate social responsibility reports by market-listed companies: contents and their evolution. *International Journal of Innovation and Sustainable Development* 7 (1): 1–26.
- Frost, G., and N. Martinov-Bennie. 2010. *Sustainability reporting assurance: market trends and information content*. Melbourne: CPA Australia.
- Früh, W. 1994. *Realitätsvermittlung durch Massenmedien*. Wiesbaden: Westdeutscher Verlag.
- Géczy, P. 2014. Big data characteristics. *The Macrotheme Review* 3 (6): 94–104.
- George, G., M. Haas, and A. Pentland. 2014. Big data and management. *Academy of Management Journal* 57 (2): 321.
- GRI. 2016. GRI G4 - Sustainability reporting guidelines. <https://www2.globalreporting.org/standards/g4/Pages/default.aspx>. Accessed 4 Mar 2017.
- Gürtürk, A., and R. Hahn. 2016. An empirical assessment of assurance statements in sustainability reports: smoke screens or enlightening information? *Journal of Cleaner Production* 136: 30–41.
- Guthrie, J., and F. Farneti. 2008. GRI sustainability reporting by Australian public sector organizations. *Public Money & Management* 28 (6): 361–366.
- Habek, P., and R. Wolniak. 2016. Assessing the quality of corporate social responsibility reports: the case of reporting practices in selected European Union member states. *Quality and Quantity* 50 (1): 399–420.
- Habermas, J. 1984. *The theory of communicative action*, vol. 1. Boston: Beacon Press.
- Hansmann, K., and C. Ringle. 2005. Strategische Erfolgswirkung einer Teilnahme an Unternehmensnetzwerken: eine empirische Untersuchung. *Die Unternehmung—Swiss Journal of Business Research and Practice* 59 (3): 217–236.
- Heinrich, B., and M. Klier. 2015. Datenqualitätsmetriken für ein ökonomisch orientiertes Qualitätsmanagement. In *Daten- und Informationsqualität—Auf dem Weg zur Information Excellence*, ed. K. Hildebrand, M. Gebauer, H. Hinrichs, and M. Mielke, 49–68. Wiesbaden: Vieweg + Teubner.
- Hinrichs, H. 2002. *Datenqualitätsmanagement in data warehouse-systemen*. Oldenburg: Universität Oldenburg.
- Hodge, K., N. Subramaniam, and J. Stewart. 2009. Assurance of sustainability reports: Impact on report users' confidence and perceptions of information credibility. *Australian Accounting Review* 19 (3): 178–194.
- Hsueh, J. 2018. Governance structure and the credibility gap: experimental evidence on family businesses' sustainability reporting. *Journal of Business Ethics* 153 (2): 547–568.
- Ioannou, I., and G. Serafeim. 2014. The consequences of mandatory corporate sustainability reporting: evidence from four countries. *Harvard Business School Working Paper*, pp. 11–100.
- ISO, International Organization for Standardization. 2015. Quality management systems—requirements. *DS/EN ISO 9001*: 2015.

- ISO, International Organization for Standardization. 2010. Guidance on Social Responsibility. *ISO 26000*: 2010.
- Kepler, J., V. Gadepally, P. Michaleas, N. Schear, M. Varia, A. Yerukhimovich, and R. Cunningham. 2014. Computing on masked data: a high performance method for improving big data veracity. *High Performance Extreme Computing Conference (HPEC) 2014* IEEE, pp. 1–6.
- Kleine, A. 1995. *Entscheidungstheoretische Aspekte der Principal-Agent-Theorie*. Heidelberg: Physica Verlag.
- Knebel, S., and P. Seele. 2015. Quo vadis GRI? A (critical) assessment of GRI 3.1 A + non-financial reports and implications for credibility and standardization. *Corporate Communications An International Journal* 20 (2): 196–212.
- Köhnen, G. 1990. *Glaubwürdigkeit. Untersuchungen zu einem psychologischen Konstrukt*. München: Psychologie Verlags Union.
- KPMG. 2013. The KPMG survey of corporate responsibility reporting 2013. <https://assets.kpmg.com/content/dam/kpmg/pdf/2015/08/kpmg-survey-of-corporate-responsibility-reporting-2013.pdf>. Accessed 4 Mar 2017.
- Krcmar, H. 2015. *Informationsmanagement*. Berlin: Springer Gabler Verlag.
- Kühnle, S., and A. Dingelstedt. 2014. Kausalität. In *Handbuch Methoden der empirischen Sozialforschung*, ed. N. Baur and J. Blasius, 1017–1028. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lehrer, Christiane, Alexander Wieneke, Jan vom Brocke, Reinhard Jung, and Stefan Seidel. 2018. How big data analytics enables service innovation: materiality, affordance, and the individualization of service. *Journal of Management Information Systems* 35 (2): 424–460.
- Liljenström, H., and U. Svedin. 2005. System features, dynamics and resilience—some introductory remarks. In *Micro meso macro—addressing complex systems couplings*, ed. H. Liljenström and U. Svedin, 1–18. Singapore: World Scientific Publishing.
- Lock, I. 2016. Glaubwürdigkeit in der CSR-Kommunikation—Entwicklung eines legitimitätsbasierten Ansatzes. *Publizistik* 61: 413–429.
- Lock, I., and P. Seele. 2016. The credibility of CSR (corporate social responsibility) reports in Europe. *Journal of Cleaner Production* 122: 186–200.
- Lukoianova, T., and V.L. Rubin. 2014. Veracity roadmap: Is big data objective, truthful and credible? *Advances In Classification Research Online* 24 (1): 4–15.
- Manetti, G. 2011. The quality of stakeholder engagement in sustainability reporting: empirical evidence and critical points. *Corporate Social Responsibility and Environmental Management* 18: 110–122.
- Manetti, G., and L. Becatti. 2009. Assurance services for sustainability reports: standards and empirical evidence. *Journal of Business Ethics* 87 (1): 289–298.
- Manetti, G., and S. Toccafondi. 2012. The role of stakeholders in sustainability reporting assurance. *Journal of Business Ethics* 107 (3): 363–377.
- Michnik, J., and M.-C. Lo. 2009. The assessment of the information quality with the aid of multiple criteria analysis. *European Journal of Operational Research* 195: 850–856.
- Milne, M., and R. Gray. 2013. W(h)ither ecology? The triple bottom line, the global reporting initiative, and corporate sustainability reporting. *Journal of Business Ethics* 118: 13–29.
- Mishra, S., and S. Modi. 2013. Positive and negative corporate social responsibility, financial leverage, and idiosyncratic risk. *Journal of Business Ethics* 117 (2): 431–448.
- Morhardt, J., S. Baird, and K. Freeman. 2002. Scoring corporate environmental and sustainability reports using GRI 2000, ISO 14031 and other criteria. *Corporate Social Responsibility and Environmental Management* 9: 215–233.
- Müller, Oliver, Maria Fay, and Jan vom Brocke. 2018. The effect of big data and analytics on firm performance: an econometric analysis considering industry characteristics. *Journal of Management Information Systems* 35 (2): 488–509.
- Nachhaltigkeitskodex, Deutscher. 2016. Nachhaltigkeitskodex erfüllt zukünftige EU-Berichtspflichten. <http://www.deutscher-nachhaltigkeitskodex.de/de/dnk/eu-berichtspflicht.html>. Accessed 04 Mar 2017.
- Natarajan, Prashant, John C. Frenzel, and Detlev H. Smaltz. 2017. *Demystifying big data and machine learning for healthcare*. Boca Raton: CRC Press.
- Nawratil, U. 2006. *Glaubwürdigkeit in der sozialen Kommunikation*. München: Westdeutscher Verlag.
- Nawratil, U. 1999. Glaubwürdigkeit als Faktor im Prozess medialer Kommunikation. In *Glaubwürdigkeit im Internet—Fragestellung, Modelle, empirische Befunde*, ed. P. Rössler, and W. Wirth, 15–31. München: Fischer.

- O'Dwyer, B., D. Owen, and J. Unerman. 2011. Seeking legitimacy for new assurance forms: the case of assurance on sustainability reporting. *Accounting, Organizations and Society* 31 (1): 31–52.
- Perego, P., and A. Kolk. 2012. Multinationals' accountability on sustainability: the evolution of third-party assurance of sustainability reports. *Journal of Business Ethics* 110 (2): 173–190.
- Perrini, F. 2006. The practitioner's perspective on non-financial reporting. *California Management Review* 48: 73–103.
- Pratt, J., and R. Zeckhauser. 1985. *Principals and agents: the structure of business*. Boston: Harvard Business School Press.
- Przyborski, A., and M. Wohlrab-Sahr. 2014. Forschungsdesigns für die qualitative Sozialforschung. In *Handbuch Methoden der empirischen Sozialforschung*, ed. N. Baur and J. Blasius, 117–133. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Robinson, M., A. Kleffner, and S. Bertels. 2011. Signaling sustainability leadership: empirical evidence of the value of DJSI membership. *Journal of Business Ethics* 101 (3): 493–505.
- Rohweder, J., G. Kasten, D. Malzahn, A. Piro, and J. Schmid. 2015. Informationsqualität—Definitionen, Dimensionen und Begriffe. In *Daten- und Informationsqualität—Auf dem Weg zur Information Excellence*, ed. K. Hildebrand, M. Gebauer, H. Hinrichs, and M. Mielke, 25–46. Wiesbaden: Vieweg + Teubner.
- Rossiter, J. 2002. The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing* 19: 305–335.
- Schaltegger, S. 1997. Information costs, quality of information and stakeholder involvement. *Corporate Social Responsibility and Environmental Management* 4 (11): 87–97.
- Schön, D. 2016. *Planung und Reporting—Grundlagen, Business Intelligence, Mobile BI und Big-Data-Analytics*. Wiesbaden: Springer Verlag.
- Sethi, P., T. Martell, and M. Demir. 2015. Enhancing the role and effectiveness of corporate social responsibility (CSR) reports: the missing element of content verification and integrity assurance. *Journal of Business Ethics* 4: 1–24.
- Seufert, A. 2016. Die Digitalisierung als Herausforderung für Unternehmen: Status Quo, Chancen und Herausforderungen im Umfeld BI & Big Data. In *Big Data—Grundlagen, Systeme und Nutzungspotenziale*, ed. D. Fasel and A. Meier, 39–58. Wiesbaden: Springer Vieweg.
- Shankaranarayan, G., M. Ziad, and R. Wang. 2003. Managing data quality in dynamic decision environments: an information product approach. *Journal of Database Management* 14 (4): 14–32.
- Simnett, R., A. Vanstraelen, and W. Chua. 2009. Assurance on sustainability reports: an international comparison. *The Accounting Review* 84 (3): 937–967.
- Spelthahn, S., F. Livia, and U. Demele. 2009. Glaubwürdigkeit in der Nachhaltigkeitsberichterstattung. *uwf UmweltWirtschaftsForum* 17 (1): 61–68.
- Suchman, M.C. 1995. Managing legitimacy: strategic and institutional approaches. *Academy of Management Review* 20: 571–610.
- Thomson, I., and J. Bebbington. 2005. Social and environmental reporting in the UK: a pedagogic evaluation. *Critical Perspectives on Accounting* 16: 507–533.
- Tole, A. 2013. Big data challenges. *Database Systems Journal* 4 (3): 31–40.
- Trommsdorff, V. 1975. *Die Messung von Produktimages für das Marketing: Grundlagen und Operationalisierung*. Köln: Heymann Verlag.
- Trotman, A., and K. Trotman. 2015. Internal audit's role in GHG emission and energy reporting: evidence from audit committees, senior accountants and internal auditors. *Auditing A Journal of Practice and Theory* 34 (1): 199–230.
- Van Altena, A., P. Moerland, A. Zwinderman, and S. Olabarriaga. 2016. Understanding big data themes from scientific biomedical literature through topic modeling. *Journal of Big Data* 3 (23): 1–21.
- Vargas-Solar, G., J. Espinosa-Oviedo, and L. Zechinelli-Martini. 2016. Big continuous data: dealing with velocity by composing event streams. In *Big data concepts, theories, and applications*, ed. S. Yu and G. Song, 1–28. Cham: Springer Verlag.
- Vormedal, I., and A. Ruud. 2009. Sustainability reporting in Norway—an assessment of performance in the context of legal demands and socio-political drivers. *Business Strategy and the Environment* 18 (4): 207–222.
- Wallage, P. 2000. Assurance on sustainability reporting: an auditor's view. *Auditing A Journal of Practice and Theory* 19 (1): 53–65.
- Wang, R., and D. Strong. 1996. Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems* 12 (4): 5–33.

- WCED, World Commission on Environmental Development. 1987. *Our common future*. Oxford: Oxford University Press.
- Weber, T. 2014. Nachhaltigkeitsberichterstattung als Bestandteil marketingbasierter CSR-Kommunikation. In *CSR und Reporting—Nachhaltigkeits- und CSR-Berichterstattung verstehen und erfolgreich umsetzen*, ed. M. Fifka, 95–106. Berlin: Springer Gabler.
- Weiber, R., and D. Mühlhaus. 2014. *Strukturgleichungsmodellierung*. Berlin: Springer.
- Windolph, S. 2011. Assessing corporate sustainability through ratings: challenges and their causes. *Journal of Environmental Sustainability* 1 (1): 61–80.
- Wirth, W. 1999. Methodologische und konzeptionelle Aspekte der Glaubwürdigkeitsforschung. In *Glaubwürdigkeit im Internet. Fragestellungen, Modelle, empirische Befunde*, eds. P. Rössler, and W. Wirth. München: Reinhard Fischer Verlag.
- Wolling, J. 2004. Qualitätserwartungen, Qualitätswahrnehmungen und die Nutzung von Fernsehserien. *Publizistik* 49 (2): 171–193.
- Ylijoki, O., and J. Porras. 2016. Perspectives to definition of big data: a mapping study and discussion. *Journal of Innovation Management* 4 (1): 69–91.
- Zimmermann, F. 2016. Was ist Nachhaltigkeit—eine Perspektivenfrage? In *Nachhaltigkeit wofür? Von Chancen und Herausforderungen für eine nachhaltige Zukunft*, ed. F. Zimmermann, 1–24. Berlin: Springer Spektrum.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.