Diginomics Working Paper

Research Group Diginomics
Digitalization - Economy - Society

Social media analytics in operations and supply chain management: Opportunities, challenges and paradoxes

Aseem Kinra, Fabian Siekmann, Herbert Kotzab

June 2022 No. 0018

Universität Bremen

Impressum:

Diginomics Working Paper

ISSN: 2701-6307

DOI: https://doi.org/10.26092/elib/1621

Published by:
Universität Bremen, Diginomics Research Group, Max-von-Laue-Straße 1, 28359 Bremen, Germany

Editor:
Prof. Dr. Lars Hornuf
Phone: +49(0)421 218 66820
E-mail: hornuf@uni-bremen.de

https://www.uni-bremen.de/graduiertengruppe-diginomics/forschung/diginomics-working-paper

This working paper will appear in L. Hornuf (Ed.), *Diginomics Research Perspectives: The Role of Digitalization in Business and Society*. Cham: Springer International Publishing.

Social media analytics in operations and supply chain management: Opportunities, challenges and paradoxes

Aseem Kinra, Fabian Siekmann and Herbert Kotzab

Industrial and academic communities in the field of operations and supply chain management (OSCM) have been paying increasing attention to social media analytics (SMA). However, the disparity of social media has inspired new ways of thinking about how data are produced, organized and analyzed. This chapter addresses how OSCM is affected by this disparity and provides an overview of SMA use, applications and challenges in the domain. A directed content analysis of current, application-oriented research is carried out to review SMA in OSCM from a signaling theory perspective. In particular, we shed light on data sources, opportunities, challenges, paradoxes and current managerial issues and seek to inform research practices and policy in order to advance operations and supply chain management research. The chapter contributes to the understanding of SMA in OSCM by identifying a set of paradoxes and challenges that have not previously been identified in OSCM research. By relating SMA to social media data sources and OSCM activities, it shed light on preferred sources and application scenarios and discusses the imponderables of social media signal processing in OSCM.

10.1 Introduction

Operations and supply chain management (OSCM) is the management of processes by which products and services are designed, procured, produced and delivered. It includes transformational activities to create products and services, as well as the management of materials, money, people and information (Seyedghorban et al., 2021). Spearman and Hopp (2020) define operations as "an act that utilizes resources to transform one or more attributes of an entity or set of entities into some good or service that is required to satisfy some external demand," encompassing a wide range of human activities. Mentzer and colleagues (2008) portray the logistics, marketing and production function of the firm as the domain for operations management, while the supply chain management domain crosses organizational boundaries. Considering the operations area as part of the supply chain, the enormous complexity of operations and supply chain management becomes apparent. The decision scope of managers changes as it expands across functional and organizational boundaries, and suitable decision-making tools and frameworks are ultimately needed to enhance organizational performance.

Through advancements in digital technologies and data analysis techniques, big data analytics provides the possibility to positively affect the operation and supply chain design process and operational decision-making (Waller and Fawcett, 2013). Today, only 5% of all existing data is in such structured forms as spreadsheets or relational databases. Unstructured data, such as text, images, audio and video sometimes lack the structural organization required by machines for analysis, but constitute 95% of big data (Gandomi and Haider, 2015). In practice, 80% of all data possessed by organizations are in the form of unstructured textual data (Wenzel and Van Quaquebeke, 2018) and the expected increase in the use of image and video media will proliferate unstructured data yet more in the future. This abundance of unstructured big data holds enormous potential for OSCM.

Organizations, and all the operations within them, are part of the social world (Spearman and Hopp, 2020). Since operations and supply chain activities are ultimately run by people for people, they can benefit from the exchange of textual information publicly available in the form of user-generated content on social media (Kinra et al., 2019). These vast, largely untapped textual data sets have the potential to help address the management of goods and information flows under uncertainty and the dynamics of the environment (Hsuan et al., 2015; Seyedghorban et al., 2021) to advance decision support through SMA. Wamba and colleagues (2016) state that, when different types of analyses are properly applied, SMA can deliver business value to firms. However, despite the euphoria with which big data technology finds its way into organizations, research is still needed to address the many challenges and paradoxes in its adoption and use. As an example, Gupta and George (2016) describe the big data productivity paradox, referring to a failure in establishing a positive relationship between social media analytics investments and firm productivity. Similarly, George and colleagues (2014) emphasize the lack of clarity about how products, services and data can be mixed into sustainable and economical models.

The purpose of this chapter is to depict current and future perspectives on the use of textual social media analytics in OSCM. It aims to provide an up-to-date assessment of the current state of social media analytics in operations and supply chain management, and to identify the problems that can arise from SMA-based decision-making in OSCM. Signaling theory as well as related organizational theories are adopted as a lens to understand the applications, problems and paradoxes related to the usage of SMA in OSCM. A directed content analysis (Krippendorff, 2004) of SMA in OSCM research is carried out to describe the current developments in the growing field of SMA, highlighting its opportunities and challenges. Rather than a thorough analysis, this chapter seeks to provide a representative overview of the evolving research domain.

The chapter is structured as follows. First, a basic understanding of the terminology in the domain is established, and an overview of the current research on unstructured social media data and its big data characteristics is provided. Signaling theory is then applied to shed light on the process of social media communication between consumers and the organization. Various data sources for SMA are classified and presented along with exemplary text analytics techniques and application scenarios to illustrate the current state of research. This is followed by a detailed exploration of the problems, challenges and a sampling of the paradoxes that organizations may be faced with in their future adoption and use of SMA. Examples are provided of the pressing issues that are necessitated by the use of big data and SMA in practice. The chapter concludes with a summary of the main findings.

10.2 Defining Big Data and SMA in OSCM

Big data is an umbrella term often used to describe data that is massive, complex and generated in real-time, therefore requiring sophisticated management and analytical and processing techniques to extract management insights (Waller and Fawcett, 2013; Dubey et al., 2019). According to Wenzel and van Quaquebeke (2018, p. 3), “Big Data can be defined as observational records that may be exceptionally numerous, highly heterogeneous, and/or generated at high rate and systematically captured, aggregated, and analyzed to useful ends.”

In recent years, BDA has increasingly found its way into the research field of OSCM (Rahimi et al., 2020).The analysis of big data is known as big data analytics (BDA), also referred to as advanced analytics, which describes the collection of data, analytical tools, computer algorithms and techniques to derive meaningful insights and patterns from the collected large data sets (Kamble and Gunasekaran, 2020) using pattern recognition techniques, statistics, machine learning, artificial intelligence and data mining (Abbott, 2014; Kinra et al., 2019; Fieberg et al., 2022). One of the most important categories of big data is social media, which, taking a broad understanding of the term, refers to various communication technology platforms that allow people to share information and opinions connecting both existing and potential customers. Social media data falls under the definition of big data as it is generated in large quantities and in real time, can take the form of network graphs, text, image and video data and the associated metadata, and often has a high degree of uncertainty due to the low level of traceability. The analysis of the data generated on social media is referred to as SMA. As a comprehensive concept, SMA encompasses BDA, social network analysis, topic modeling and sentiment analysis (Kamble and Gunasekaran, 2020). The SMA process is conceptualized along three categories: aggregation, analysis and interpretation of social media data (Wang et al., 2020). Accordingly, companies that successfully apply SMA have the ability to collect, store and monitor social media data relevant to OSCM, analyze the collected data, and interpret and utilize the results to support OSCM decision-making (Huang et al., 2019). Although significant overlaps in the use of the different concepts may exist, Table 1 presents some selected definitions about the concepts and aids in differentiating these from each other as far as possible.

Concept	**Source**	**Definition**
Big Data	Wenzel & van Quaquebeke (2018)	Big data can be defined as observational records that may be exceptionally numerous, highly heterogeneous, and/or generated at a high rate and systematically captured, aggregated, and analyzed to useful ends.
	Gartner IT-Glossary (2021)	Big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision-making.
	Jeble et al. (2018)	Big data is defined as datasets that are too large for traditional data processing systems and therefore require new technologies to handle them.
	Dubey et al. (2019)	The term "big data" is often used to describe massive, complex and real-time data that requires sophisticated management, analytical and processing techniques to extract management insights.
Big Data Analytics (BDA)	Abbott (2014)	BDA is defined as the process of discovering meaningful patterns of data using pattern recognition techniques, statistics, machine learning, artificial intelligence and data mining.
	Gandomi & Heider (2015)	Big data analytics can be viewed as a sub-process in the overall process of "insight extraction" from big data.
	Kamble & Gunasekaran (2020)	Big data analytics (BDA) defined as collection of data, analytical tools, computer algorithms and techniques to derive meaningful insights and patterns from the collected large data sets.

Social Media Analytics (SMA)	Chen et al. (2012)	A method to uncover what customers think and feel by analyzing structured and unstructured online data dispersed across a vast array of online sources.
	Zeng et al. (2010)	Informatics tools and frameworks to collect, monitor, analyze, summarize and visualize social media data to facilitate conversations and interactions to extract useful patterns and intelligence.
	Fan & Gordon (2014)	SMA as an interdisciplinary modelling and analytical paradigm consisting of three steps: (1) capturing data from various sources; (2) understanding data using various analytics and models; and (3) summarizing and presenting the findings for decision-making.
	Wamba et al. (2016)	SMA is a more comprehensive analysis tool than BDA or social network or sentiment analysis, as SMA can encompass BDA, social network analysis and sentiment analysis to understand consumers.

***Table 10.1** Conceptual orientation*

Another commonly accepted definition for Big Data is proposed by Gartner, defining big data according to its 3Vs: "high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" (Pettey, 2017). Veracity is mentioned in the literature as a fourth V, reflecting the (un)certainty of the available data (Fosso Wamba et al., 2015). Together these also represent the different characteristics of big data and need to be addressed within the all forms and categories, including SMA. However, before we perform a more in-depth analysis, we turn our attention to the main reasons and the dominant theoretical rationale for SMA in OSCM, namely signals and signaling theory. This will aid us in the overall analysis of the state of the art in the field, and the main challenges and paradoxes in the use of SMA.

10.3 Some theoretical perspectives on textual social media data for OSCM

10.3.1 User-generated content as signals for OSCM

The markets in which the organizations operate are increasingly influenced by competition, technological change, shorter product life-cycles and increasingly demanding customers, who expect responsiveness to a dynamic set of requirements, which challenges firms in most industries (Bernardes and Hanna, 2009). The organization's responsibility is therefore to understand external stakeholders such as customers and suppliers and adjust its objectives accordingly (Slack et al., 2010). Since OSCM is concerned with producing goods and services and meeting the requirements of customers, their perceptions are of great interest to operations management (Slack et al., 2010), whose goal is to serve demand as effectively and efficiently as possible. Social media generates huge amounts of information on companies and customers in real-time. This is because social media enables customers to share their experiences of products and services with a broad public, which not only consists of potential consumers, but also the manufacturers or distributors of these products. These publicly expressed customer experiences possess the properties of *signals*, which are "activities or attributes of individuals in a market which, by design or accident, alter the beliefs of, or convey information to, other individuals in the market" (Spence, 1974).

10.3.2 Signaling theory and the social media environment

Signaling theory is essentially concerned with reducing information asymmetry between two parties (Spence, 1974, 2002). Originating in the overcoming of information asymmetries in labor markets, signaling theory has been applied in numerous management research fields, expanding the range of potential signals and the contexts in which signaling occurs (Connelly et al., 2011; Wallenburg et al., 2017; Banerjee et al., 2020). The key components of signaling theory are signalers, signals, receivers and feedback, all operating in a signaling environment.

Applied to the present example of social media as a signaling environment, this means that the customer as signaler sends information via social media to a known or unknown receiver – potentially the focal firm or future consumers – which then processes the signal to receive information. The receiver's reaction can be considered as feedback that is sent to the signaler. The process is visualized in Figure 1. Signaling theory proposes that to reduce the uncertainty of social media information, receivers will screen signalers with a focus on the observability, usefulness and trustworthiness of their signals (Banerjee et al., 2020). Despite social media communication being clearly observable and proven to be useful for organizations, the trustworthiness of the signals and their senders remains uncertain, as receivers (organizations) do not exercise control over crowds (Lukyanenko et al., 2019). In addition, the evaluation of the signals is subject to an often an individualized process of interpretation (Branzei et al., 2004). The trustworthiness of these signals, and thus of the data from which information is generated for OSCM, is accordingly assessed on the basis of the data context and the expertise of the interpreter (Banerjee et al.,

2020). Both understanding the cultural context in social media and domain expertise are therefore crucial to overcoming the veracity problem associated with social media data uncertainty for OSCM, which is discussed in greater detail below in the context of SMA challenges. First, we highlight the main opportunities and applications of SMA in OSCM in the next section.

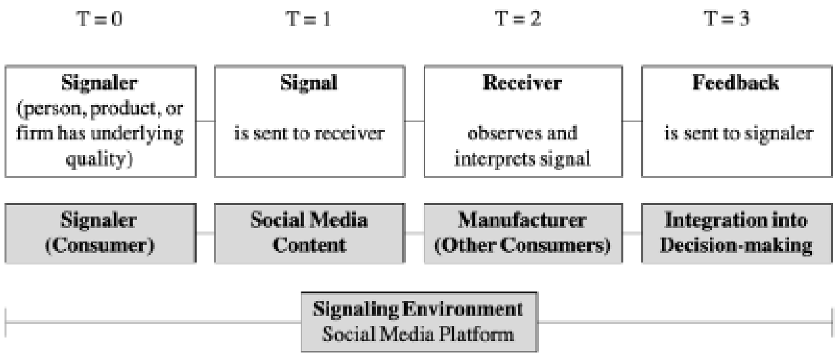

Fig. 10.1 The consumer – manufacturer signaling process in the social media environment (based on Connelly et al., 2011)

10.4. SMA in OSCM: The main data sources, opportunities and applications

In this section we make use of the supply chain operations reference (SCOR) model (APICS Supply Chain Council, 2017) to illustrate: (1) how social media data can be used in OSCM for the main operational activities of planning, sourcing, making, delivery and return; (2) the specific data sources that have been documented in the literature; and (3) how these can be used to enhance operational performance along the main performance dimensions of costs, time, quality, flexibility and innovativeness (Kamble and Gunasekaran, 2020).

10.4.1 Big data sources and social media

Social media encompasses various internet-based applications that allow content generation, sharing and continuous modification by users in a collaborative fashion. In line with the broader definition applied earlier in this chapter, the content may include written messages, as well as pictures, videos, animation, audio materials and web links to external sites, each constituting the interface between customers and the organization (Tóth et al., 2019).

As indicated above, social media is only one of many types of big data. The different data sources that are used by organizations to capture value from big data can be mapped by differentiating internal and external data sources. Internal data feeds from either existing data or self-generated data that may be crowd-sourced or tracked, generated or collected in a similar fashion. External data sources are populated by acquired data, customer-provided data or freely available data. Freely available data might be open data, web-crawled data or social media data, and the influential potential of this data for decision-making is shown by Hartmann and colleagues (2016), who state that nine out of ten organizations rely on external data sources. Figure 2 presents a taxonomy of SMA data sources for OSCM, categorizing social media as an external and freely available big data source and thus differentiating it from internal data sources and acquired or customer-provided data.

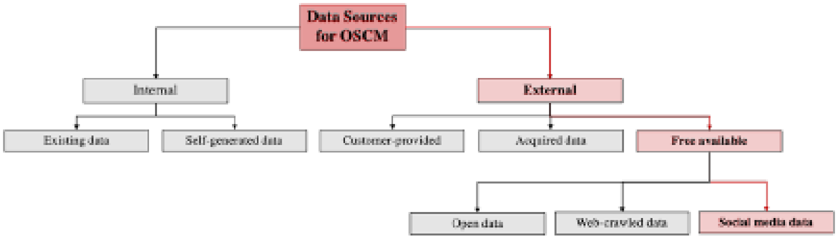

Fig. 10.1 Taxonomy of data sources for OSCM (based on Hartmann et al., 2014)

At present, unstructured data analysis methods are typically applied to the use of textual data (Adnan and Akbar, 2019a). This chapter thus forgoes analysis of image and video data, and instead focuses on the examination of text data, whose availability is a precondition of SMA. Social media data and web-crawled data are hereafter discussed from this point of view.

10.4.2 Different sources for SMA and how these have been covered by the OSCM literature

Kinra and colleagues (2019) differentiate between traditional textual data sources and social media-based sources of textual data. While traditional textual data sources such as research articles, newspapers and magazines, as well as company, consulting and government reports, provide information to the group of readers, they lack an interactive "social" element present in contemporary social media. Social media can be classified into traditional social networking sites (e.g., Facebook, Instagram), web-based communities (e.g., brand communities), collaborative projects (e.g., Wikipedia), user-generated content communities (e.g., blogs), multimedia sharing sites (e.g., YouTube or TikTok) and virtual worlds (e.g., Second Life or the proposed Metaverse). Mobile social media include microblogs such as Twitter, and internal networks (Cheng and Krumwiede, 2018) such as WhatsApp and Telegram that allow for continuous mobile communication.

Applying a broader perspective, Abrahams and colleagues (2012) define social media as online services that provide for decentralized, user-level content creation, social interaction and open membership, including discussion forums, public wikis, open online communities (social networks) and public groups, as well as customer product reviews, public visitor comments and user-contributed news articles. We adopt this definition in order to provide an overview of SMA usage in the current

OSCM literature. Our broad analysis of the literature shows that tweets and product reviews are the most frequently used SMA data sources. Tweets, for example, are often used as a source of transportation information, as they are produced at high speed and thus make it possible to obtain information in near-real time (Chen et al., 2016; Gu, et al., 2016; Zavala and Ramirez-Marquez, 2019). With the exception of delivery, product reviews are used in all SCOR activities. OSCM makes use of the highly relevant data directly related to the product (Mudambi and Schuff, 2010), which enables a detailed understanding of the customer's perspective on the product or service (Min et al., 2018). It is striking that analysis of data not only focuses on the data itself, but also frequently considers the associated structured metadata. Examples are given by Gu and colleagues (2016), who use location data to map traffic incidents, and Mudambi and Schuff (2010), who incorporate the star ratings into the understanding of reviews of search and experience goods. A similar picture emerges when looking at the use of blogs and forum posts. Beheshti-Kashi and colleagues (2019) and Zavala and Ramirez-Marquez (2019) show how the combination of text and its metadata can be used for generating value-added information both for real-time decision-making and for longer time frames.

10.4.3 SMA applications according to the SCOR model

SMA is utilized by organizations in connection-building with both the upstream- and downstream supply chain, covering both the business-to-customer and business-to business markets. It supports the sharing of information, assisting the flow of information between customers, departments and persons, creating interorganizational networks (Devi and Ganguly, 2021). On the basis of the resulting information flows, the functions of social media for OSCM can be categorized as company–public information exchange, inter-organizational information exchange and intra-organizational information exchange. These flows enable efficient communication and effectiveness, but also increased environmental awareness (Huang et al., 2019). Operational efficiency benefits from enhanced communication, where social media data is used as information technology for profit maximization or cost reduction, helping firms to survive in competitive environments (Moyano-Fuentes et al., 2021). As an interface to the organization's environment, social media also serves as a sensing opportunity that enables higher situational awareness of companies. The improved coordination enabled by such access to additional data is shown to increase operational performance (Devi and Ganguly, 2021). These two basic functions – as connecting tool and data source – are exemplified in the SCOR activities detailed below.

10.4.3.1 Planning: Demand forecasting and inventory management

Demand forecasting is the predominant application within SCOR activity planning, and benefits from using operations information, accounting variables and market data for prediction (Devi and Ganguly, 2021). Sodero and Rabinovich (2017) develop a demand-forecasting model integrating the sentiments of discussion forum contributors into decision-making. Wood and colleagues (2016) aim to integrate sentiment analysis of product reviews into upstream firms' demand-sensing activities. Hou and colleagues (2017) and Chong and colleagues (2016) showcase the predicting power of sentiment in combination with quantitative review characteristics, while Lee and colleagues (2017) shed light on the issue of sentiment manipulation on Twitter. Papanagnou and Matthews-Amune (2018) combine sales data with unstructured multiple social media sources to improve inventory management. Yuan and colleagues (2018) investigate daily sentiments in online reviews through a joint sentiment–topic model approach.

In the fashion industry context, Beheshti-Kashi and colleagues (2019) demonstrate how to forecast fashion trends based on blog-posts, and Choi (2018) develops a model to illustrate demand confidence of managers in the presence of social media. Irani and colleagues (2017) study the capabilities needed for collaborative co-design and collaborative design-to-order at the interface of joint product development with potentially disruptive effects on strategic inventory control. On the basis of the examples listed, it can be concluded that SMA can inform the planning process and reduce uncertainty in decision-making for operations planning.

10.4.3.2 Sourcing: Supplier evaluation, information sharing and idea capturing

Sourcing is a SCOR activity that has received comparatively little research attention (Huang et al., 2019). Mishra and Singh (2018) use twitter to identify and trace back product quality issues along the supply chain. Lin and colleagues (2018) demonstrate a supplier evaluation model integrating customer preferences in Google search and on Twitter. Grant and Preston (2019) illustrate that social influence plays a powerful role in supporting knowledge-sharing with suppliers. Jiang and colleagues (2017) propose a method to locate quality-related reviews corresponding to different aspects of product quality and production systems at the interface of the making activity of the SCOR model. These examples show that SMA can be used to track the supply chain and evaluate how consumer requirements are being integrated even in the upstream supply chain.

10.4.3.3 Making: Product development and risk management

The making activities of product development and business risk management are further OSCM areas that benefit from both supplier and customer involvement in new product design (Devi and Ganguly, 2021). Ramanathan and colleagues (2017) present a model for integrating review data to understand customer satisfaction. Cheng and Krumwiede (2018) use social media data from internal social networks to enhance supplier involvement in new product development. Sigala (2014) suggests considering customer input in the tourism sector, and Guo and colleagues

(2015) identify critical factors in attraction management based on online reviews. These use cases suggest that decision-making in product and service development can benefit from SMA results through a better understanding of market needs.

10.4.3.4 Delivery: Transportation information and customer service

Gu and colleagues (2016) and Liu and Qu (2016) present a methodology to crawl, process and filter tweets in real time to detect traffic incidents, and Albuquerque and colleagues (2016) use tweets from authorities in charge of road networks or by news agencies to identify traffic-related events. Han and colleagues (2019) combine social media and RFDI data to harness the power of crowdsourcing and the Internet of Things in disaster response, and Chen and colleagues (2016) examine the discovery and utilization of relevant tweets for real-time disaster management. Bhattacharjya and colleagues (2016) analyze tweets related to logistics services for purposes of identifying effective and ineffective social media customer service strategies. Gkiotsalitis and Stathopoulos (2016) investigate social media users' willingness to travel in relation to varying daytimes, and Cottrill and colleagues (2017) investigate information management strategies during major events. Such decision-making benefits from coverage of the transport network by the crowd and from the metadata enabling localization to ensure optimal transportation and customer service.

10.4.3.4.5 Return: Returns forecasting

Minnema and colleagues (2016) examine the impact of intrinsic review characteristics such as valence, volume, and variance on return decisions, finding that overly positive reviews induce more purchases, but also more returns. Sahoo and colleagues (2018) investigate how product reviews reduce uncertainty and the corresponding probability of returns among risk-averse, rational consumers. Zavala and Ramirez-Marquez (2019) propose a visualization framework to detect product disruptions based on customer sentiment from Twitter and Facebook to reduce response time. These examples show how operations management can use consumer input to obtain detailed information about product shortcomings and return behavior and thus decide on possible countermeasures at an early stage.

Table 2 presents examples where SMA is found in the OSCM literature (organized according to the SCOR activities of planning, sourcing, making, delivery and return) and shows which data sources are used for different application fields. Our review reveals that planning and delivery are the SCOR activities in which SMA is used more frequently, with reviews being used strikingly more often for the former and Twitter data for the latter.

The next section delves into the various problems and challenges presented by the use of SMA in OSCM. The main challenges for the utilization of SMA can be associated with the various SMA techniques, big data characteristics, and the paradoxes that the use of big data poses for the effective deployment of SMA for OSCM.

Contribution	Data Source	SCOR Activity	Application
Sodero & Rabinovich (2017)	forum posts	planning	demand forecasting
Wood et al. (2016)	reviews		demand forecasting
Hou et al. (2017)	reviews		demand forecasting
Chong et al. (2016)	reviews		demand forecasting
Papanagou & Matthew (2018)	multiple		demand forecasting
Yuan et al. (2018)	reviews		demand forecasting
Swain & Cao (2018)	multiple		demand forecasting
Beheshti-Kashi et al. (2019)	blogs		demand forecasting
Lee et al. (2017)	reviews		demand shaping
Choi (2018)	general		inventory management
Irani et al. (2017)	multiple		inventory management
Mishra & Singh (2018)	tweets	sourcing	waste management
Grant & Preston (2019)	general		knowledge sharing
Lin et al. (2018)	tweets, Google		supplier evaluation
Markova & Petkovska (2013)	general		supplier evaluation
Jiang et al. (2017)	reviews		supplier evaluation
Ramanathan et al. (2017)	reviews	making	product development
Cheng & Krumwiede (2018)	internal net-works		product development
Sigala (2014)	reviews, general		customer involvement
Guo et al. (2016)	reviews		business risk management
Singh et al. (2017)	tweets	delivery	traffic monitoring
Albuquerque et al. (2016)	tweets		traffic monitoring
Chen et al. (2016)	tweets		traffic monitoring
Gkiotsalitis & Stathopoulos (2016)	tweets		traffic monitoring
Liu & Qu (2016)	internal net-works		traffic monitoring
Han et al. (2019)	general, RFID		traffic monitoring
Cottrill et al. (2017)	tweets		traffic monitoring
Gu et al. (2016)	tweets		traffic monitoring
Bhattachajya et al. (2016)	tweets		customer service
Minnema et al. (2016)	reviews	return	returns forecasting
Sahoo et al. (2018)	reviews		returns forecasting
Zavala & Ramirez-Marquez (2019)	forum posts		returns forecasting

Table 10.2 A selection of applications for SMA in OSCM

10.5 SMA in OSCM: Challenges and Characteristics

10.5.1 Signal processing: Some current challenges within the dominant SMA techniques in OSCM

Recent advances in computer-aided natural language processing have led to a wide range of techniques for analyzing social media text. Topic modeling and sentiment analysis are two dominant techniques which are applied frequently in OSCM. Although these are not the only techniques and numerous other concepts for text processing have found their way into the OSCM literature, topic modeling and sentiment analysis are particularly illustrative of some of the difficulties that may arise in algorithm-driven language processing.

Topic modeling can be described as a statistical model that identifies topics as latent constructs that span a particular text corpus. The technique is typically used for finding clusters of words to identify major concepts or trends within and across corpuses or to reduce their dimensionality (Kinra et al., 2019). With regard to the OSCM domain, topic modeling has been applied to theorize the field and to reveal insights that might otherwise be lost in the abundance of data (Chae, 2015; Bansal et al., 2020). More practically, Yuan and colleagues (2018) extract latent information on product features to predict sales performance, benefiting from simplified corpus preparation compared to supervised models. Kinra and colleagues (2019) apply topic modeling to newspaper articles to gain insights on the adoption of driverless cars without the need of potentially biased data. Focusing on comparably long and targeted newspaper documents as data inputs, the authors avoid the common limitations associated with document length and topic quality (Albalawi, Yeap and Benyoucef, 2020).

Sentiment analysis is a similarly popular classification technique that assigns documents to sentiment categories, which are typically labeled as positive, neutral or negative. Sentiment analysis relies on the input of keywords, topics and character sequences for categorization. To reduce the amount of work required to generate this input data, standard corpora are frequently applied in OSCM for sentiment analysis (Kinra et al., 2019). Swain and Cao (2019) show how sentiment analysis in a supply chain context can be applied to unstructured social media documents to extract sentiments and opinions of customers, suppliers and employees, drawing conclusions on supply chain performance. However, the limitations of the method go beyond the technical, as the true polarity of a document is influenced by contextual features such as negotiation, syntax and modality (Abrahams et al., 2012). This becomes clear if we look at the word "good" in the comment "this design is not good". Although the supposedly positive word "good" is included in this sentence, an inference about its real sentiment depends upon the context of its use, which in this case is negative.

Even if advances in language processing have navigated simple pitfalls such as the example given, it is assumed that the understanding of complex contextdepending language patterns, such as humor and irony, will remain a weak point in these techniques for the time being. To complicate matters, sentiment-indicative words may differ across domains, limiting the use of standard text corpora for sentiment analysis (Loughran and McDonald, 2011).

10.5.2 Dealing with big data characteristics in SMA for effective signal capture

The characteristics of big data – the four Vs outlined by Gartner (2017) – have implications for the use of SMA for OSCM. At first glance, the high volume of data and the velocity with which it is produced, as well as its variety and potentially unsettled veracity, represent enormous potential for comprehensively improving customer understanding and reducing uncertainty (Fosso Wamba et al., 2015; Hazen et al., 2018). The four Vs nevertheless pose distinct challenges for OSCM compared to handling of traditional, structured data.

10.5.2.1 Volume

Built into the very notion of big data, volume poses a problem and is influenced by numerous factors. For example, advances in storage technology imply that data volumes that are now considered big data may not cease to expand in the future. The measurement of data volumes can also be misleading because, for example, two data sets of the same size may require different processing, and volume is therefore contextual. As a result, it is practically impossible to determine boundaries for what constitutes big data (Gandomi and Haider, 2015).

The practical handling of the volume in social media data also poses information technology challenges for OSCM. Parssian and colleagues (2004) highlight that data quality issues increase with the amounts of data collected, as implementing corresponding processes is costly in terms of software and associated human resources. Lukyanenko and colleagues (2019) state that little is known about data capture in the context of unanticipated phenomena or how to foster discoveries from user-generated content in the abundance of available data. Gu and colleagues (2016) point out that, in the context of real-time traffic incident detection, the volume of information that can be generated from tweets depends on the number of users and their spatial distribution.

The handling of high social media data volumes also poses a variety of cognitive challenges for OSCM. For example, excessive social media connectivity is argued to lead to emotional exhaustion, impairing turnover intentions of supply chain professionals (Tang et al., 2019). Exploring the question of how the volume of review data affects consumer behavior, Minnema and colleagues (2016) find that volume affects the purchase decision, but has little or no effect on product returns.

10.5.2.2 Velocity

The proliferation of mobile devices has led to an increase in the amount of data generated, particularly on social media, which requires real-time analysis to obtain the most current information possible. Given the limitations of traditional data management systems, big data technologies are needed to instantaneously process this continuously produced data (Gandomi and Haider, 2015). Even given the technical challenges of processing the data flows in real time, such processing requires compromises and hence may not be very precise. The respective tools are often designed without any hypotheses, favoring correlations over causation, resulting in susceptibility to errors (Scholz, 2017).

The OSCM literature provides various examples of how to manage the velocity at which social media data is generated. Kirac and Milburn (2018) illustrate the tradeoff between accuracy and timeliness of social media analysis and show the strong potential for social data integration to improve disaster response operations. Liu and Qu (2016) utilize crowdsourced textual data from social media platforms to characterize the dynamics and uncertainty of road conditions and minimize collective travel time of vehicles. Gu and colleagues (2016) crawl, process and filter tweets from Twitter in real time to detect traffic incidents. However, the trade-offs necessary for real-time data processing along supply chains result in a new type of uncertainty related to the data itself, as the velocity comes at the cost of accuracy. At a certain point, the efficiency of the tool ceases to compensate for the unreliability of its results, resulting in what Lechler and colleagues (2019) refer to as an "uncertainty dilemma."

10.5.2.3 Variety

The variety of big data is believed to benefit the granularity and accuracy of decision-making due to the variety of sources that can be explored (McAfee and Brynjolfsson, 2012). Unstructured data present on social media nevertheless bring enormous challenges as they lack schema or standardization and come in multiple formats (Adnan and Akbar, 2019b). Recent technological advances allow companies to use the various types of unstructured big data (Kinra et al., 2020). However, the range of types and sources is broad, as data in social media may take the form of text, images, audio or video (Adnan and Akbar, 2019a), and methods such as image-based analytics are still in their infancy (Gandomi and Haider, 2015).

While numerous examples exist for textual SMA, the analysis of natural language is far from trivial. Papanagnou and Matthews-Amune (2018) use a combination of Google search, YouTube and newspaper articles as data sources for predicting demand for medicines, highlighting the advantages of inclusion of different data structures. Challenges arising from different textual data sources are also illustrated by Zavala and Ramirez-Marquez (2019), who collect Facebook data when identifying product recalls, but base their analysis on supposedly unbiased Twitter data. Kinra and colleagues (2019) state the need to examine data sources individually with an eye to their limitations and barriers.

10.5.2.4 Veracity

Due to the variety of big data sources and the accessibility that characterizes social media, the assurance of trustworthiness of data is a major challenge for OSCM (Wilkin et al., 2020), particularly in ensuring valid and relevant insights (Gandomi and Haider, 2015). Although larger text corpora are believed to reduce factors restricting data veracity, decision-makers need to have confidence in the generated results (Kinra et al., 2020).

In addressing data veracity issues, Zavala and Ramirez-Marquez (2019) use Twitter instead of Facebook data, referring to possible biases and errors caused by bots. Research also shows that the data generated via social media is prone to manipulation. For example, Lee and colleagues (2017) find evidence of sentiment manipulation on Twitter through organizations. Wu and colleagues (2020) state that studies focus on organizations as contributors of fake reviews while neglecting the role of users and the platforms themselves. Decision-makers are therefore challenged to validate the results derived from SMA to reduce stakeholder-related veracity issues (Bansal et al., 2020).

10.6 Coping with big data paradoxes in order to overcome signal distortion in OSCM

The access among organizations and practitioners of OSCM to the socio-technical ecosystem of social media presents decision-makers with enormous challenges associated with the use of consumer data. The visibility of the signalers and the receiver's analytic capabilities require balanced decision-making to avoid generating long-term competitive disadvantages through opportunities that are beneficial to the operations function in the short term.

10.6.1 Connectivity paradox

The connectivity paradox is described as a situation when one party fears being disconnected from the other party and starts to use technologies that establish a connection that is so close, the actor in fear of being disconnected also needs to take heed of practices to disconnect. Efforts to strengthen customer loyalty may result in information being provided to the operations function, but the company's long-term goals may be jeopardized by a dysfunctional relationship (Leonardi and Treem, 2020). The connectivity paradox concerns the point in the signaling process where the signaler decides to send the signal but cannot control the recipients of the signals due to the public nature of the social media platform as a signaling environment. Considering the number of ways in which social media data is already being analyzed, it is easy to forget that customers are already deliberately avoiding the production of usable data.

Consequently, OSCM is faced with the challenge of balancing the extent to which the skimming of data produces added value or facilitates countermeasures and whether more connectivity ultimately leads to better outcomes. An example of such negative implications of social media usage is given by Tang and colleagues (2019), who demonstrate a negative effect of social media connectivity on turnover intentions of supply chain professionals. As a consequence, social media users may reduce the observability of their behavior, thus consciously limiting a necessary characteristic of signals to be examined.

10.6.2 Performance paradox

As described above, SMA requires a great deal of effort to identify relevant data, clean it, and prepare the results. Therefore, much of the effort is invisible and the information generated is neither standardized nor explicit. This is one basis for the performance paradox, wherein actors expend so many resources to make task performance visible that this communication effort diminishes actual performance (Leonardi and Treem, 2020).

From the signaling perspective, the performance paradox is located between $T = 1$ and $T = 2$, when the receiver (i.e., the organization) observes and interprets the transmitted signals. Due to the lack of standard practices in BDA and their domain-knowledge dependability (Waller and Fawcett, 2013), disturbances may occur in the dissemination of information generated from SMA. The performance paradox thus confronts OSCM with the task of not only considering the work itself and its underlying technology in order to measure performance, but also keeping track of the extent to which the results are tangible for others than the analysts.

With regard to the use of SMA in companies, this means that a certain understanding of the technology's results should be established in order to create a target-oriented balance between the workload and the communication effort. A key task is therefore to integrate the technology into established corporate processes (Spearman and Hopp, 2020). Hazen and colleagues (2014) detail approaches for combining the measurement of data quality with established methods in the OM field, such as control charts. Zavala and Ramirez-Marquez (2019) show how visualization frameworks can be used to prepare insights from Twitter data in an intelligible way.

10.6.3 Transparency paradox

The transparency paradox describes the situation where organizations put effort into increasing transparency, but in doing so produce a volume and variety of communication and information that obstruct a clear view of their analysis processes. A high degree of transparency does not reduce secrecy, but counterintuitively increases it, since a flood of information that is difficult to process reduces visibility (Hald and Kinra, 2019). While a high degree of transparency in business processes is socially considered desirable, it allows companies to conceal aspects of their operations (Leonardi and Treem, 2020).

The transparency paradox relates to T = 3 of the signaling process, when the receiver sends feedback to the signaler. Connelly and colleagues (2011) emphasize the importance of this feedback in providing information to the signalers about the effectiveness of their signals in order to reduce information asymmetry. However, in the case of social media communication, the organization is not necessarily the primary addressee of the signal, hence it faces the difficult task of creating transparency without being able to address the entirety of the signalers in response to the respective signals. Due to the variety of touchpoints between OSCM practitioners and consumers via social media, the challenge lies in providing an appropriate degree of visibility of relevant and critical analysis processes. Accordingly, the need for communication relates not only to internal stakeholders, but also to the signal providers themselves. Zavala and Ramirez-Marquez (2019) take steps toward resolving such issues by applying a visual analytics framework based on established control charts to support real-time decision-making with low-threshold, complementary, relevant information.

10.6.4 Power paradox

Information processing along the phases identified by signaling theory shows that the entire process of SMA is accompanied by paradoxes, which are formed from the SMA-related integration of the customer's social context into operations management. The paradoxes occur when the SMA-enabled opportunity for a supposedly better customer understanding leads to actual deterioration. The connectivity paradox is particularly relevant in the interaction between signaler (e.g., customer) and receiver (i.e., observing organization). The performance paradox affects the interface between receiver and feedback (i.e., the integration of knowledge into business processes), while the transparency paradox affects the entire process from signal transmission to information integration. As these paradoxes reveal, the analysis of social media data creates winners and losers. Without a properly defined framework, both operations management and customers, with their respective needs, find themselves in an uncertain state that is not beneficial to either party – what is known as the power-paradox (Richards and King, 2013).

Although the subject needs further research, the different paradoxes are understood as emerging from and impacting the signaling process as visualized in Figure 3. If a consensus cannot be reached on which analysis methods are appropriate, then the full potential of SMA will not be realized. The challenge is to strike a balance that allows operations management to add value without compromising the self-determination of the customer. Consequently, it is necessary to consider holistically what level of visibility is desired, and where restrictions may be needed to meet the needs of stakeholders.

The next section covers examples of real-life breakdowns that may occur in OSCM as a result of the challenges and paradoxes in the use of social media and SMA that have been discussed above.

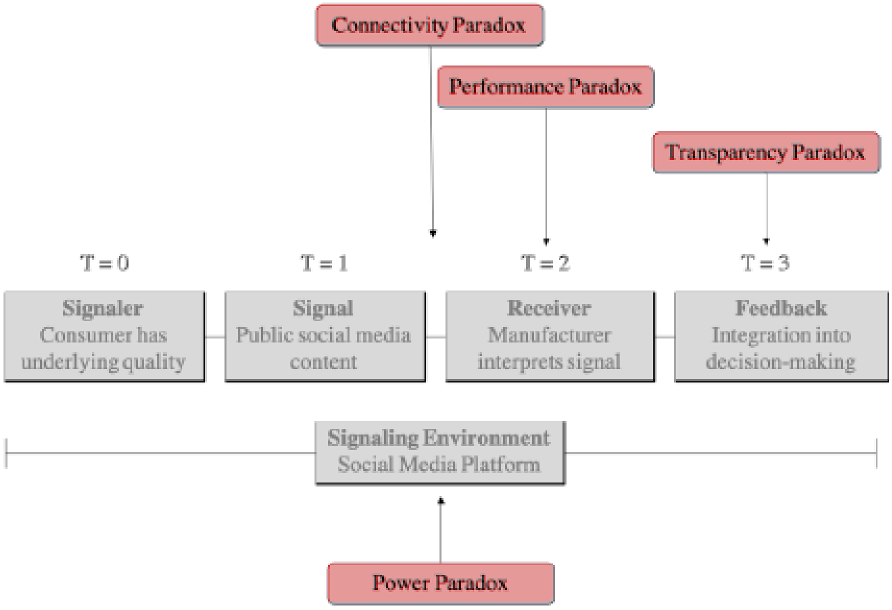

***Fig. 10.2** SMA paradoxes in the signaling process*

10.7 Current issues in the practice of SMA in OSCM

10.7.1 Security and data breaches

OSCM decisions are increasingly becoming data-driven. The sections above show why data is considered a key resource for the company and why companies invest in protecting it. This is reflected in corporate spending. Gartner expects annual expenditure in the area of IT security to increase by around 10% (Pettey, 2017). However, cybersecurity is characterized by turbulence and change and is further complicated by the need to manage the multiple tiers in modern supply chains (Melnyk et al., 2021). Despite increases in spending, data breaches continue to occur. A data breach is "the unauthorized acquisition of data in any form or format containing sensitive information that compromises the security or confidentiality of such information and creates a reasonable risk of misuse" (Cannon and Kessler, 2007), putting customers at risk, incurring considerable costs to organizations, and threatening business reputations (Syed, 2019).

Chatterjee and Sokol (2019) differentiate between three types of data breaches: *data disclosure* is the unauthorized access to confidential data by attackers; *data alteration* refers to the manipulation, alteration, corruption and deletion of data; and *data denial* refers to the situation where data access is suppressed by attackers. The recent example of the A.P. Moller–Maersk cyberattack shows that organizations and supply chains from all industries are exposed to the risk of data breaches (Ritchie, 2019). As a result of a malware attack in 2017, the shipping giant suffered a failure of its IT systems, affecting a diversity of its business units, including container shipping, port and tug-boat operations, drilling services, oil and gas production, and transportation. While 95% of stock carried in containers by Maersk arrived on time, the disruption necessitated enormous personnel efforts on the part of the company to deal with the consequences of the incident. Further illustration of potential data security problems can be seen with UPS and FedEx (whose data breaches leaked the information of thousands of customers; FedEx, 2017; Bisson, 2020), Total Quality Logistics (where carrier information was exposed due to a cyberattack; Forde, 2020), and Forward Air (which warned of revenue losses following a ransomware attack; Tabak, 2020).

The gateways for such risks are many and various and are based on technological as well as procedural and behavioral weaknesses, making them difficult to protect against. To deal with increasingly frequent malware and ransomware attacks (e.g., WannaCry, which infected 230,000 computers worldwide in 2017), organizations need to realize that such attacks represent a fundamental business problem, potentially affecting operations along the whole supply chain. Consumers are also becoming increasingly deliberate in what types of data they share, with media coverage of data breaches increasingly raising awareness of the risks involved. Such wide awareness nevertheless allowed Maersk to enhance its reputation as a trustworthy brand by openly communicating about the incident and its effects on its own systems. This example illustrates how, despite the business challenge, secure data handling and privacy can become a differentiator and even a source of competitive advantage (Kaplan et al., 2020).

10.7.2 Moral and ethical challenges

SMA is the automation of human decision-making by social media-based algorithms that produce probability-aided information. However, the information generated from this must inevitably be regarded as uncertain and is thus often unsuitable for triggering action. This is particularly true for unstructured data and user-generated content from social media. It is important to include this limitation when basing decisions on SMA-generated insights. Given the lack of transparency of algorithms, especially in the area of deep learning, the link between data and the information generated by the algorithm is often not traceable. In practice, this leads to problems when it comes to accountability for decisions, limiting the action-initiation ability of SMA insights (Mittelstadt et al., 2016). Opening this black box is critical in all analytics systems where supervised or unsupervised machine learning is used (Goebel et al., 2018; Kinra et al., 2019; Fieberg et al., 2022). As a remedy, explainable artificial intelligence approaches have been developed to increase interpretability of such decision support systems (Guidotti et al., 2018) and have begun to find their way into the OSCM literature (e.g. Senoner et a., 2021).

SMA algorithms are prone to the "garbage in, garbage out" principle. SMA-based insights can therefore only be as reliable as the input they process, and the open character of social media platforms can have serious ramifications for this reliability. As an example, Lee and colleagues (2017) find instances of sentiment manipulation in movie reviews and thus illustrate that the processing of social media data should not be considered infallible. It is also necessary to question the extent to which the recommendations for action derived from the insights can be reconciled with the company's long-term objectives (Mittelstadt et al., 2016). For example, while it may make sense to adjust prices according to the situation or customer group from an operational point of view, this may run counter to the organization's marketing strategy and thus reduce the long-term performance of the company and the entire supply chain. A practical example of this is provided by Ketzenberg and colleagues (2020), who clarify that identifying abusive potential customer returns is not sufficient, as this may jeopardize relationships with profitable customers.

These are only some of the examples that constitute the debate on the moral and ethical challenges of SMA in OSCM. The analysis of Mittelstadt and colleagues (2016) is recommended for further discussion of such issues.

10.8 Conclusion

This chapter provides a summary of current and future perspectives on the use of textual social media analytics in OSCM, with the aim of presenting the current state of the art, and identifying its opportunities and challenges. The detailed literature analysis demonstrates that the use of SMA is still in its infancy. This chapter also details the use of product reviews and Twitter content as the dominant SMA data sources. Activities in planning and delivery operations in particular benefit from the analysis of this type of data, while sourcing, making and return operations activities have received comparatively little attention in the research literature to date.

To shed light on the problems that can arise in the use of SMA for decision-making in OSCM, the existing research reveals the characteristics of big data that hinder effective social media data processing. A set of paradoxes, the connectivity-performance-transparency-power paradoxes, are identified alongside the challenges they pose to OSCM, as well as how SMA activities may run counterproductive to their promised information gains. The impact of cybercrime as part of responsible management for the use of potentially sensitive consumer data is also discussed, as well as the moral and ethical challenges that organizations and managers face in advancing their applications of SMA.

The chapter is also one of the first studies to apply signaling theory as a lens to examine the use of social media data processing in OSCM. By identifying the inherent paradoxes it uniquely illustrates the behavioral visibility challenges for the signaling process in operations, and postulates why SMA may not have unfolded in OSCM as speculated by Waller and Fawcett (2013). The main implications of the study are that practitioners should note the shortcomings of analyzing textual social media data and act accordingly. Knowledge of these difficulties can serve as a guideline in the integration of SMA into business practices. Ideally, managers should try to bundle competencies along the interface between data science and OSCM. For future research directions, the study reveals the need to explore both the possible applications of the technology across the range OSCM activities, as well as the reasonable depth of the insights gained from SMA.

To holistically reflect the state of research, the period and scope for articles included in future studies should be regularly adjusted to match the changing dynamics of the research field. Based on the research presented here, it is clear that the potential of SMA analytics is distributed across the entire breadth of the research field and that the challenges are as diverse as the data itself. The current state of research suggests that scholars should also seek to develop a deeper understanding of the data being analyzed, the patterns of behavior on social media, and the underlying analytics methods in order to fully realize the potential of SMA. From this perspective, more in-depth research is warranted into the identified paradoxes within OSCM. If research succeeds in making this interdisciplinary leap, this promising field can be expected to bear exciting and actionable results in both research and practice.

References

Abbott, D. (2014). *Applied predictive analytics: Principles and techniques for the professional data analyst*. John Wiley & Sons, Inc.

Abrahams, A. S., Jiao, J., Wang, G. A., & Fan, W. (2012). Vehicle defect discovery from social media. *Decision Support Systems*, 54(1), 87–97. doi: 10.1016/j.dss.2012.04.005.

Adnan, K., & Akbar, R. (2019a). An analytical study of information extraction from unstructured and multidimensional big data. *Journal of Big Data*. Springer International Publishing. doi: 10.1186/s40537-019-0254-8.

Adnan, K., & Akbar, R. (2019b). Limitations of information extraction methods and techniques for heterogeneous unstructured big data. *International Journal of Engineering Business Management*, 11, 1–23. doi: 10.1177/1847979019890771.

Albalawi, R., Yeap, T. H., & Benyoucef, M. (2020). Using topic modeling methods for short-text data: A comparative analysis. *Frontiers in Artificial Intelligence*, 3(July), 1–14. doi: 10.3389/frai.2020.00042.

Albuquerque, F. C., Casanova, M. A., Lopes, H., Redlich, L. R., De Macedo, J. A. F., Lemos, M., De Carvalho, M. T. M., & Renso, C. (2016). A methodology for traffic-related Twitter messages interpretation. *Computers in Industry*, 78, 57–69. doi: 10.1016/j.compind.2015.10.005.

APICS Supply Chain Council (2017). *The supply chain operations reference model (SCOR)*. Available at: http://www.apics.org/docs/default-source/scor-p-toolkits/apics-scc-scor-quick-reference-guide.pdf.

Banerjee, A., Ries, J. M., & Wiertz, C. (2020). The impact of social media signals on supplier selection: Insights from two experiments. *International Journal of Operations and Production Management*, 40(5), 531–552. doi: 10.1108/IJOPM-05-2019-0413.

Bansal, P., Gualandris, J., & Kim, N. (2020). Theorizing supply chains with qualitative big data and topic modeling. *Journal of Supply Chain Management*, 56(2), 7–18. doi: 10.1111/jscm.12224.

Beheshti-Kashi, S., Pannek, J., & Kinra, A. (2019). Complementing decision support and forecasting risk in supply chain with unstructured data. *IFAC-PapersOnLine*, 52(13), 1721–1726. doi: 10.1016/j.ifacol.2019.11.449.

Bernardes, E. S., & Hanna, M. D. (2009). A theoretical review of flexibility, agility and responsiveness in the operations management literature: Toward a conceptual definition of customer responsiveness. *International Journal of Operations and Production Management*, 29(1), 30–53. doi: 10.1108/01443570910925352.

Bhattacharjya, J., Ellison, A., & Tripathi, S. (2016). An exploration of logistics related customer service provision on Twitter: The case of e-retailers. *International Journal of Physical Distribution & Logistics Management*, 46(6).

Bisson, D. (2020). *UPS says phishing incident might have exposed some customers' data*. Available at: https://securityboulevard.com/2020/01/ups-says-phishing-incident-might-have-exposed-some-customers-data/.

Branzei, O., Ursacki-Bryant, T. J., Vertinsky, I., & Zhang, W. (2004). The formation of green strategies in chinese firms: Matching corporate environmental responses and individual principles. *Strategic Management Journal*, 25(11), 1075–1095. doi: 10.1002/smj.409.

Cannon, D. M., & Kessler, L. (2007). Danger: Corporate data breach! *The Journal of Corporate Accounting & Finance*, 18(5), 41–49. doi: 10.1002/jcaf.

Chae, B. (2015). Insights from hashtag #supplychain and Twitter analytics: Considering Twitter and Twitter data for supply chain practice and research. *International Journal of Production Economics*, 165, 247–259. doi: 10.1016/j.ijpe.2014.12.037.

Chatterjee, C., & Sokol, D. D. (2019). Data security, data breaches, and compliance. In C. Chatterjee, D. D. & Sokol (Eds.), *Cambridge handbook of compliance* (pp. 936–948). Cambridge University Press.

Chen, X., Elmes, G., Ye, X., & Chang, J. (2016). Implementing a real-time Twitter-based system for resource dispatch in disaster management. *GeoJournal*, 81(6), 863–873. doi: 10.1007/s10708-016-9745-8.

Cheng, C. C. J., & Krumwiede, D. (2018). Enhancing the performance of supplier involvement in new product development: The enabling roles of social media and firm capabilities. *Supply Chain Management*, 23(3), 171–187. doi: 10.1108/SCM-07-2017-0230.

Choi, T. M. (2018). Incorporating social media observations and bounded rationality into fashion quick response supply chains in the big data era. *Transportation Research Part E: Logistics and Transportation Review*, 114, 386–397. doi: 10.1016/j.tre.2016.11.006.

Chong, A. Y. L., Li, B., Ngai, E. W., Ch'ng, E., & Lee, F. (2016). Predicting online product sales via online reviews, sentiments, and promotion strategies: A big data architecture and neural network approach. *International Journal of Operations & Production Management*.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, 37(1), 39–67. doi: 10.1177/0149206310388419.

Cottrill, C., Gault, P., Yeboah, G., Nelson, J. D., Anable, J., & Budd, T. (2017). Tweeting transit: An examination of social media strategies for transport information management during a large event. *Transportation Research Part C: Emerging Technologies*, 77, 421–432. doi: 10.1016/j.trc.2017.02.008.

Devi, Y., & Ganguly, K. (2021). Social media in operations and supply chain management: A systematic literature review to explore the future. *Operations and Supply Chain Management: An International Journal*, 14(2), 232–248. doi: 10.31387/oscm0450299.

Dubey, R., Gunasekaran, A., Childe, S. J., Blome, C., & Papadopoulos, T. (2019). Big data and predictive analytics and manufacturing performance: Integrating institutional theory, resource-based view and big data culture. *British Journal of Management*, 30(2), 341–361. doi: 10.1111/1467-8551.12355.

FedEx (2017). FedEx files 10-K with additional disclosure on cyber-attack affecting TNT Express Systems. Available at: https://investors.fedex.com/news-and-events/investor-news/investor-news-details/2017/FedEx-Files-10-K-with-Additional-Disclosure-on-Cyber-Attack-Affecting-TNT-Express-Systems/default.aspx.

Fieberg, C., Hesse, M., Loy, T., & Metko, D. (2022). Machine Learning in Accounting Research. In L. Hornuf (Ed.), *Diginomics Research Perspectives: The Role of Digitalization in Business and Society*, (pp. @@@). Cham: Springer International Publishing.

Forde, M. (2020). TQL cyber breach is latest example of the industry's vulnerability to hacking. Available at: www.supplychaindive.com/news/tql-cyber-breach-industry-vulnerability-hacking/573174/.

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How "big data" can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165, 234–246. doi: 10.1016/j.ijpe.2014.12.031.

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

George, G., Haas, M. R., & Pentland, A. (2014). Big data and management. *Academy of Management Journal*, 57(2), 321–326. doi: 10.18352/bmgn-lchr.9536.

Gkiotsalitis, K., & Stathopoulos, A. (2016). Joint leisure travel optimization with user-generated data via perceived utility maximization. *Transportation Research Part C: Emerging Technologies*, 68, 532–548. doi: 10.1016/j.trc.2016.05.009.

Goebel, R., Chander, A., Holzinger, K., Lecue, F., Stumpf, S., Kieseberg, P., Holzinger, A., Goebel, R., Chander, A., Holzinger, K., Lecue, F., Akata, Z., Goebel, R., Chander, A., Holzinger, K., & Lecue, F. (2018). Explainable AI: the new 42? To cite this version: HAL Id: hal-01934928', 2nd International Cross-Domain Conference for Machine Learning and Knowledge Extraction (CD-MAKE).

Grant, S. B., & Preston, T. A. (2019). Using social power and influence to mobilise the supply chain into knowledge sharing: A case in insurance. *Information and Management*, 56(5), 625–639. doi: 10.1016/j.im.2018.10.004.

Gu, Y., Qian, Z., & Chen, F. (2016). From Twitter to detector: Real-time traffic incident detection using social media data. *Transportation Research Part C: Emerging Technologies*, 67, 321–342. doi: 10.1016/j.trc.2016.02.011.

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, 51(5). doi: 10.1145/3236009.

Guo, Y., Sun, S., Schuckert, M., & Law, R. (2015). Online feedback and attraction management: An exploration of the critical factors in effective operations. *Asia Pacific Journal of Tourism Research*, 21(8), 883–904. doi: 10.1080/10941665.2015.1080740.

Hald, K. S., & Kinra, A. (2019). How the blockchain enables and constrains supply chain performance. *International Journal of Physical Distribution and Logistics Management*, 49(4), 376–397. doi: 10.1108/IJPDLM-02-2019-0063.

Han, S., Huang, H., Luo, Z., & Foropon, C. (2019). Harnessing the power of crowdsourcing and Internet of Things in disaster response. *Annals of Operations Research*, 283(1–2), 1175–1190. doi: 10.1007/s10479-018-2884-1.

Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2014). Capturing value from big data: A taxonomy of data-driven business models used by start-up firms. *International Journal of Operations and Production Management*, 36(10), 1382–1406. doi: 10.1108/IJOPM-02-2014-0098.

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, 154, 72–80. doi: 10.1016/j.ijpe.2014.04.018.

Hazen, B. T., Skipper, J. B., Boone, C. A., & Hill, R. R. (2018). Back in business: Operations research in support of big data analytics for operations and supply chain management. *Annals of Operations Research*, 270(1–2), 201–211. doi: 10.1007/s10479-016-2226-0.

Hou, F., Li, B., Chong, A. Y. L., Yannopoulou, N., & Liu, M. J. (2017). Understanding and predicting what influence online product sales? A neural network approach. *Production Planning and Control*, 28(11–12), 964–975. doi: 10.1080/09537287.2017.1336791.

Hsuan, J., Skjott-Larsen, T., Kinra, A., & Kotzab, H. (2015). *Managing the Global Supply Chain* (4th ed.). Frederiksberg: CBS Press.

Huang, S., Potter, A., & Eyers, D. (2019). Social media in operations and supply chain management: State-of-the-art and research directions. *International Journal of Production Research*, 58(6), 1893–1925. doi: 10.1080/00207543.2019.1702228.

Irani, Z., Sharif, A. M., Papadopoulos, T., & Love, P. E. D. (2017). Social media and Web 2.0 for knowledge sharing in product design. *Production Planning and Control*, 28(13), 1047–1065. doi: 10.1080/09537287.2017.1329955.

Jiang, C., Liu, Y., Ding, Y., Liang, K., & Duan, R. (2017). Capturing helpful reviews from social media for product quality improvement: A multi-class classification approach. *International Journal of Production Research*, 55(12), 3528–3541. doi: 10.1080/00207543.2017.1304664.

Kamble, S. S., & Gunasekaran, A. (2020). Big data-driven supply chain performance measurement system: a review and framework for implementation. *International Journal of Production Research*, 58(1), 65–86. doi: 10.1080/00207543.2019.1630770.

Kaplan, J., Soller, H., Anant, V., & Donchak, L. (2020). *The consumer-data opportunity and the privacy imperative*. Chicago: McKinsey & Company.

Ketzenberg, M. E., Abbey, J. D., Heim, G. R., & Kumar, S. (2020). Assessing customer return behaviors through data analytics. *Journal of Operations Management*, 66(6), 622–645. doi: 10.1002/joom.1086.

Kinra, A., Hald, K. S., Mukkamala, R. R., & Vatrapu, R. (2020). An unstructured big data approach for country logistics performance assessment in global supply chains. *International Journal of Operations and Production Management*, 40(4), 439–458. doi: 10.1108/IJOPM-07-2019-0544.

Kinra, A., Kashi, S. B., Pereira, F. C., Combes, F., & Rothengatter, W. (2019). Textual data in transportation research: Techniques and opportunities. In *Mobility Patterns, Big Data and Transport Analytics* (pp. 173–197). Elsevier. doi: 10.1016/b978-0-12-812970-8.00008-7.

Kirac, E., & Milburn, A. B. (2018). A general framework for assessing the value of social data for disaster response logistics planning. *European Journal of Operational Research*, 269(2), 486–500. doi: 10.1016/j.ejor.2018.02.011.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage Publications Limited. doi: 10.2307/2288384.

Lechler, S., Canzaniello, A., Roßmann, B., von der Gracht, H. A., & Hartmann, E. (2019). Real-time data processing in supply chain management: Revealing the uncertainty dilemma. *International Journal of Physical Distribution and Logistics Management*, 49(10), 1003–1019. doi: 10.1108/IJPDLM-12-2017-0398.

Lee, S.-Y., Qiu, L., & Whinston, A. (2017). Sentiment manipulation in online platforms: An analysis of movie tweets. *Production and Operations Management*, 27(3), 393–416. doi: 10.1111/ijlh.12426.

Leonardi, P. M., & Treem, J. W. (2020). Behavioral visibility: A new paradigm for organization studies in the age of digitization, digitalization, and datafication. *Organization Studies*, 41(12), 1601–1625. doi: 10.1177/0170840620970728.

Lin, K. P., Hung, K. C., Lin, Y. T., & Hsieh, Y. H. (2017). Green suppliers performance evaluation in belt and road using fuzzy weighted average with social media information. *Sustainability (Switzerland)*, 10(1). doi: 10.3390/su10010005.

Liu, S., & Qu, Q. (2016). Dynamic collective routing using crowdsourcing data. *Transportation Research Part B: Methodological*, 93, 450–469. doi: 10.1016/j.trb.2016.08.005.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance*, 66(1), 35–65. doi: 10.2469/dig.v41.n2.20.

Lukyanenko, R., Parsons, J., Wiersma, Y. F., & Maddah, M. (2019). Expecting the unexpected: Effects of data collection design choices on the quality of crowdsourced user-generated content. *MIS Quarterly: Management Information Systems*, 43(2), 634–647. doi: 10.25300/MISQ/2019/14439.

McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review*, 90(10), 60–66.

Melnyk, S. A., Schoenherr, T., Speier-pero, C., Peters, C., Jeff, F., Friday, D., Melnyk, S. A., Schoenherr, T., & Speier-pero, C. (2021). New challenges in supply chain management : Cybersecurity across the supply chain chain. *International Journal of Production Research*, 1-22. doi: 10.1080/00207543.2021.1984606.

Mentzer, J. T., Stank, T. P., & Esper, T. L. (2008). Supply chain management and its relationship to logistics, marketing, production, and operations management. *Journal of Business Logistics*, 29(1), 31–46. doi: 10.1002/j.2158-1592.2008.tb00067.x.

Min, H., Yun, J., & Geum, Y. (2018). Analyzing dynamic change in customer requirements: An approach using review-based Kano analysis. *Sustainability (Switzerland)*, 10(3). doi: 10.3390/su10030746.

Minnema, A., Bijmolt, T. H. A., Gensler, S., & Wiesel, T. (2016). To keep or not to keep: Effects of online customer reviews on product returns. *Journal of Retailing*, 92(3), 253–267. doi: 10.1016/j.jretai.2016.03.001.

Mishra, N., & Singh, A. (2018). Use of twitter data for waste minimisation in beef supply chain. *Annals of Operations Research*, 270(1–2), 337–359. doi: 10.1007/s10479-016-2303-4.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data and Society*, 3(2), 1–21. doi: 10.1177/2053951716679679.

Moyano-Fuentes, J., Maqueira-Marín, J. M., Martínez-Jurado, P. J., & Sacristán-Díaz, M. (2021). Extending lean management along the supply chain: Impact on efficiency. *Journal of Manufacturing Technology Management*, 32(1), 63–84. doi: 10.1108/JMTM-10-2019-0388.

Mudambi, S. M., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly: Management Information Systems*, 34(1), 185–200.

Papanagnou, C. I., & Matthews-Amune, O. (2018). Coping with demand volatility in retail pharmacies with the aid of big data exploration. *Computers and Operations Research*, 98, 343–354. doi: 10.1016/j.cor.2017.08.009.

Parssian, A., Sarkar, S., & Jacob, V. S. (2004). Assessing data quality for information products: Impact of selection, projection, and cartesian product. *Management Science*, 50(7), 967–982. doi: 10.1287/mnsc.1040.0237.

Pettey, C. (2017). *5 reasons you're failing with social analytics: Drive stronger performance on your most critical priorities*. Available at: https://www.gartner.com/smarterwithgartner/5-reasons-youre-failing-with-social-analytics.

Rahimi, I., Gandomi, A. H., Fong, S. J., & Ülkü, M. A. (2020). *Big data analytics in supply chain management: Theory and applications*. Boca Raton, FL: CRC Press.

Ramanathan, U., Subramanian, N., & Parrott, G. (2017). Role of social media in retail network operations and marketing to enhance customer satisfaction. *International Journal of Operations and Production Management*, 37(1), 105–123. doi: 10.1108/IJOPM-03-2015-0153.

Richards, N. M., & King, J. H. (2013). Three paradoxes of big data. *Standford Law Review*, 66, 41–46.

Ritchie, R. (2019). *Maersk: Springing back from a catastrophic cyber-attack*, *I-Cio*. Available at: hhttps://www.i-cio.com/management/insight/item/maersk-springing-back-from-a-catastrophic-cyber-attack.

Sahoo, N., Dellarocas, C., & Srinivasan, S. (2018). The impact of online product reviews on product returns. *Information Systems Research*, 29(3), 723–738. doi: 10.1287/isre.2017.0736.

Scholz, T. M. (2017). *Big data in organizations and the role of human resource management*, *big data in organizations and the role of human resource* *management*. Frankfurt: Peter Lang International Academic Publishers. doi: 10.3726/b10907.

Senoner, J., Netland, T., & Feuerriegel, S. (2021). Using explainable artificial intelligence to improve process quality: Evidence from semiconductor fabrication. *Management Science*, forthcoming.

Seyedghorban, Z., Samson, D., & Swink, M. (2021). Quo vadis OSCM? An analysis of past and future trends in operations and supply chain management research. *Decision Sciences*, (April 2020), 1–23. doi: 10.1111/deci.12519.

Sigala, M. (2014). Customer involvement in sustainable supply chain management: A research framework and implications in tourism. *Cornell Hospitality Quarterly*, 55(1), 76–88. doi: 10.1177/1938965513504030.

Singh, A., Shukla, N., & Mishra, N. (2017). Social media data analytics to improve supply chain management in food industries. *Transportation Research Part E: Logistics and Transportation Review*, 114, 398–415. doi: 10.1016/j.tre.2017.05.008.

Slack, N., Chambers, S., & Johnston, R. (2010). *Operations management* (6th ed.). Harlow, UK: Pearson Education.

Sodero, A. C., & Rabinovich, E. (2017). Demand and revenue management of deteriorating inventory on the internet: An empirical study of flash sales markets. *Journal of Business Logistics*, 38(3), 170–183. doi: 10.1111/jbl.12157.

Spearman, M. L., & Hopp, W. J. (2020). The case for a unified science of operations. *Production and Operations Management*, 30(3), 802–814. doi: 10.1111/poms.13318.

Spence, M. (1974). *Market signaling, information transfer in hiring and related processes*. Cambridge, MA: Harvard University Press.

Spence, M. (2002). Signaling in retrospect and the informational structure of markets. *American Economic Review*, 92(3), 434–459. doi: 10.1257/00028280260136200.

Swain, A. K., & Cao, R. Q. (2019). Using sentiment analysis to improve supply chain intelligence. *Information Systems Frontiers*, 21(2), 469–484. doi: 10.1007/s10796-017-9762-2.

Syed, R. (2019). Enterprise reputation threats on social media: A case of data breach framing. *Journal of Strategic Information Systems*, 28(3), 257–274. doi: 10.1016/j.jsis.2018.12.001.

Tabak, N. (2020). *Forward Air reveals ransomware attack, warns of revenue hit*. Available at: https://www.freightwaves.com/news/news-alert-forward-air-reveals-ransomware-attack-warns-of-revenue-hit.

Tang, G., Ren, S., Chadee, D., & Yuan, S. (2019). The dark side of social media connectivity: Influence on turnover intentions of supply chain professionals. *International Journal of Operations and Production Management*, 40(5), 603–623. doi: 10.1108/IJOPM-05-2019-0391.

Tóth, Z., Liu, M., Luo, J., & Braziotis, C. (2019). The role of social media in managing supplier attractiveness: An investigation of business-to-business markets. *International Journal of Operations and Production Management*, 40(5), 625–646.

doi: 10.1108/IJOPM-04-2019-0321.

Wallenburg, C. M., Einmahl, L., Lee, K. B., & Rao, S. (2017). On packaging and product returns in online retail: Mailing boxes or sending signals? *Journal of Business Logistics*, 42(1), 291–308. doi: 10.1111/jbl.12273.

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77–84. doi: 10.1111/jbl.12010.

Wamba, S. F., Akter, S., Kang, H., Bhattacharya, M., & Upal, M. (2016). The primer of social media analytics. *Journal of Organizational and End User Computing*, 28(2), 1–12. doi: 10.4018/JOEUC.2016040101.

Wang, Y., Zhang, M., Tse, Y. K., & Chan, H. K. (2020). Unpacking the impact of social media analytics on customer satisfaction: Do external stakeholder characteristics matter? *International Journal of Operations and Production Management*, 40(5), 647–669. doi: 10.1108/IJOPM-04-2019-0331.

Wenzel, R., & Van Quaquebeke, N. (2018). The double-edged sword of big data in organizational and management research: A review of opportunities and risks. *Organizational Research Methods*, 21(3), 1–44. doi: 10.1177/1094428117718627.

Wilkin, C., Ferreira, A., Rotaru, K., & Gaerlan, L. R. (2020). Big data prioritization in SCM decision-making: Its role and performance implications. *International Journal of Accounting Information Systems*, 38, 100470. doi: 10.1016/j.accinf.2020.100470.

Wood, L. C., Reiners, T., & Srivastava, H. S. (2016). Think exogenous to excel: Alternative supply chain data to improve transparency and decisions. *International Journal of Logistics Research and Applications*, 20(5), 426–443. doi: 10.1080/13675567.2016.1267126.

Wu, Y., Ngai, E. W. T., Wu, P., & Wu, C. (2020). Fake online reviews: Literature review, synthesis, and directions for future research. *Decision Support Systems*, 132(February), p. 113280. doi: 10.1016/j.dss.2020.113280.

Yuan, H., Xu, W., Li, Q., & Lau, R. (2018). Topic sentiment mining for sales performance prediction in e-commerce. *Annals of Operations Research*, 270(1–2), 553–576. doi: 10.1007/s10479-017-2421-7.

Zavala, A., & Ramirez-Marquez, J. E. (2019). Visual analytics for identifying product disruptions and effects via social media. *International Journal of Production Economics*, 208(December), 544–559. doi: 10.1016/j.ijpe.2018.12.020.