



Investigating Relevant Data in Automotive Procurement Departments: External Shocks as Transparency Creator for Data Deficits in Decision-Making

Sven Klee¹ · Andreas Janson²

Accepted: 27 October 2023
© The Author(s) 2023

Abstract

Uncertainty becomes the new normal for organizations worldwide. Many organizations are dependent on complex global supply chains. COVID-19, but also environmental disasters or the war in Ukraine, demonstrate the volatility of supply chains. Procurement departments are the central interface between internal and external stakeholders and must manage the supply chain stability what requires fast and accurate decision-making. External shocks and sudden disruptions of central supply chains illustrated that data analytics could not prevent disruptions, although sound research on competitive advantages and numerous investments should have enabled organizations to data-driven decision-making. Rather, it became transparent, that there are numerous data deficits in organizations. We did an interview-based study with 23 procurement and supply chain experts about relevant data sets and the status of its usability. We contribute to theory and practice by uncovering relevant aspects of data and provide theoretical propositions on how decision-making can be improved in automotive procurement departments.

Keywords Data analytics · Decision-making · Procurement departments · Supply chain · Automotive manufacturers · External shocks

1 Introduction

“With uncertainty becoming the new norm for businesses, all supply chains are susceptible to disruptions” (Ambulkar et al., 2015). The outbreak of COVID-19 pandemic as a starting point of long-lasting crises, following limited productions due to the global semiconductor shortage or the current war in Ukraine are a selection of external shocks for automotive manufacturers and illustrate the frequency in recent years. Such crises also illustrate the vulnerability of supply chains and related businesses on the other,

especially within the automotive industry (Ahuja & Ngai, 2019; Gölgeci et al., 2023; Klee et al., 2023). For procurement departments, such shocks changed entire norms and ways of working (Bag et al., 2023). Therefore, supply chain resilience and necessary mechanisms for dealing with uncertainty are essential for organizations (Sengupta et al., 2022).

Procurement departments can be seen as the central link between internal and external stakeholders and, thus, as the responsible department for the entire supply chain (Pellengahr et al., 2016). Sudden and completely unprepared external shocks like the outbreak of COVID-19 demonstrated that quick but high-quality decisions are necessary. Therefore, procurement leaders are under pressure and must deal with uncertainties in their decision-making processes (Hallikas et al., 2021). Managers can make the best decisions when they are equipped with valuable data (Davenport, 2006). In today’s world, there is an infinite amount of data that can be analyzed and that can lead to important new insights (Bethaz & Cerquitelli, 2021). Several research findings show the correlation between efficient data management and competitive advantages in organizations (Chen et al., 2012; Huang et al., 2020; Olszak & Zurada, 2019; Someh et al., 2023; Popović

✉ Sven Klee
sven.klee@wi-kassel.de

Andreas Janson
andreas.janson@unisg.ch

¹ Information Systems, University of Kassel, Pfannkuchstraße 1, 34121 Kassel, Germany

² Institute of Information Management, University of St. Gallen, Müller-Friedberg-Straße 8, St. Gallen 9000, Switzerland

et al., 2018) or improved organizational ambidexterity for environmental sustainability (Shi et al., 2023). Procurement departments are often overwhelmed because necessary data sets are not properly exploited and usable or even missing (Li & Liu, 2019). As a result, such data sets lose their potential to support decision-making (Moretto et al., 2017). Data acquisition, data fusion and data-driven decision-making are generally challenging in the environment of complex supply chains (Han et al., 2021). Based on this, procurement departments need to think about necessary data sets and how they can be brought into the organization in creative ways (Hallikas et al., 2021). As an initial step, there must be an awareness of which data is relevant for specific decisions and then ways must be found to make these data sets usable. In general, two types of data sources can be distinguished. On the one hand the internal sources, like information systems, which procurement departments use, on the other hand external sources in the form of supply chain data, economic data or other public data, which are helpful for the purposes of procurement departments (Messina et al., 2020).

Our research builds on this by conducting an interview-based study with procurement experts about necessary data sets for decision-making in procurement departments and about the current usability of these data sets according to their experiences from ongoing projects or other activities in the field of data analytics. With the help of 23 interviews in 20 organizations, we want to analyze the current state in organizations and, thus, aim to use the expert knowledge for uncovering relevant aspects about necessary data sets in procurement departments. Furthermore, we want to set a focus on data usability by evaluating what data sets are already available in procurement departments, what data sets are probably available but not usable in an expedient manner and what data sets are still completely missing. In this context, we also deal with the necessary technical aspects that the interview experts told us based on their experience. The experts are procurement leaders from the automotive manufacturers, experienced consultants from projects in the procurement area with a focus on data analytics, business intelligence or related fields, as well as experts from software and data companies. Based on their experiences and knowledge, we address the following research question (RQ):

RQ: Which data sets are relevant for decision-making in automotive procurement departments and what is the status regarding the usability of these data sets?

For addressing this RQ, we start with a theoretical background for considering important concepts and evaluating related work. Building on this, we describe the conduction of our interview-based study by presenting detailed information about the interviewed experts, the interviewing process and

related analysis of our findings. A discussion of these findings includes an overview about relevant procurement data sets and a related classification of the usability of these data sets in three levels, as well as relevant insights on technical aspects as a prerequisite for the use of data. Finally, we present recommendations for action on how decision-making in automotive procurement departments can be improved with relevant data sets and the current status of usability. From a practical perspective, we aim to support procurement leaders and their departments in real situations, especially with regard to current and future crises. From a theoretical perspective, we want to use existing research findings and complement them with a look at the practical world of procurement.

2 Theoretical Background

2.1 Data as a Starting Point for Decision-Making

Data and information are essential for organizations and, therefore, part of all daily operations (Redman, 2008). A huge amount of heterogeneous data is continuously generated by humans and machines (Chiusano et al., 2021). The size of global digital data is doubling every year and is estimated to reach 175 zettabytes in 2025 (Reinzel et al., 2018). Success in the use of data depends on the quality of the input data (Bellatreche et al., 2022). There are numerous ways for defining these input data sets and their characteristics. Due to the diverse and ubiquitous amounts and types of data, appropriate data management is required (Darmont et al., 2022). In very general terms, data can be structured, semi-structured or unstructured (Kassner et al., 2015; Zrenner et al., 2017; Phillips-Wren et al., 2015). Building up on this, there are different aggregation forms like resource, database, record or item (Zrenner et al., 2017). Furthermore, data can be characterized according to their business use and have characteristics such as master data, transactional data or interactional data (Hannila et al., 2022). Based on this, an important property of data in organizations is the currency of data, for example with characteristics such as forecast, up-to-date or outdated (Zrenner et al., 2017). Complementary to this, data sources can vary widely. For example, relevant data can originate from internal systems, documents, web logs, social networks, sensors or machines (Zakir et al., 2015; Phillips-Wren et al., 2015). Such sources can be internal, external closed or external open, which is another way of data differentiation (Zrenner et al., 2017). A very different approach for characterizing data can also be a social approach. Data can be generated by an active information search, social transactions, information diffusion, social interaction or completely non-deliberate (Blazquez & Domenech, 2018).

2.2 Relevant Technical Aspects of Data for Decision-Making

If organizations want to make important, high-impact decisions based on data, there must be a focus on technical aspects, regardless of all the research approaches already discussed. Furthermore, the success strongly depends on the quality of input data and on the consideration of non-functional properties related to legal, ethical or economical aspects (Bellatreche et al., 2022). In this context, data modeling is highly relevant because it creates the basis for quality and trust. The success of data analytics depends not only on the handling of data itself, but also on the capabilities of organizations to design models for data analysis (Bellatreche et al., 2022). These data modeling skills are also paramount in topics such as artificial intelligence and machine learning, which are playing an increasingly important role in how organizations deal with data (Luckow et al., 2018). Data models in this context must be trainable, but also explainable for user acceptance and, most importantly, they must be reusable in dynamic data environments (Luckow et al., 2018). Huge amounts of data are produced, especially in manufacturing processes in the automotive industry, which can be used to, among other things, reduce warranty costs or increase customer satisfaction. Due to the volume of data, however, this also requires effective and automatic detection of meaningful information and associated data models (Leitner et al., 2014). Furthermore, sensor data from vehicles in the automotive industry represents massive potential from a data perspective. At this point, however, we are talking about petabytes of data, which illustrates the need for powerful data models and frameworks (Johanson et al., 2014). In addition to all the necessary skills and aspects required for the successful use of data, attention must also be paid to aspects that transport all data analysis into a legal and trustworthy framework. Thus, there is a necessity for privacy-enhancing technologies (PET) (Dickhaut et al., 2023; Garrido et al., 2022). With the European General Data Protection Regulation (GDPR) or the Consumer Privacy Act, official legal frameworks have already been created to ensure privacy, security, trust, and regulatory compliance (Garrido et al., 2022). These examples can only provide a small glimpse into the challenges of successfully handling data and resulting data-driven decision making. In a competitive environment, the automotive industry in particular must be responsive to changing customer wishes, market trends, and sustainability challenges (Johanson et al., 2014). In the following section, a focus is placed on the automotive supply chain and procurement area. However, the technical aspects discussed in this section are an important component of all subsequent sections.

2.3 Data Analytics in Procurement and Supply Chain Contexts

Data analytics is a set of techniques focusing on valuable and actionable insights to make smart decisions based on a large amount of data (Duan & Da Xu, 2021). Regardless of how different input data are characterized, what sources they come from, and how available they are, the reality in organizations often shows massive problems with data sets by using data analytics. Problems arise due to inconsistencies, format problems, missing definitions or even completely missing data sets (Hannila et al., 2022). For the purpose of our research, it is necessary to evaluate the most important data sets and its usability in the field of procurement. Several approaches focus on decision-making and managing uncertainty in organizations by focusing on data-driven approaches and technologies. For example, machine learning has the potential to help procurement departments manage supply chain risk by providing data-driven decision support based on algorithms (Baryannis et al., 2019). However, future research must ensure, that there are more extensive data sets that can be used for analysis (Baryannis et al., 2019). Numerous authors have explored potentials of big data and data analytics in general for mitigating risks and dealing with uncertainty in procurement departments or, more superordinate, in the entire supply chain (Vieira et al., 2019; Moretto et al., 2017; Hallikas et al., 2021). In this context, data analytics in general can also be seen as a collective term for statistical evaluations, simulations or data-based optimization models (Brintrup et al., 2020). Again, there is an important focus on historical data in order to be able to apply data analytics techniques of all kinds (Brintrup et al., 2020). In addition to data analytics, big data approaches, or the use of artificial intelligence in the procurement or supply chain context, there are other approaches that focus on the potentials offered by data. For example, secondary data can help to improve decision-making in procurement departments or in the entire supply chain context if it is available in a structured and usable form (Ellram & Tate, 2016). Secondary data can come from sources such as existing literature, census data, government information, financial data, organizational reports, or records, and are therefore data that have been produced by others (Lind et al., 2012). Another approach pursues a targeted distribution of information among actors in the supply chain (Cui & Idota, 2018). With the help of blockchain technology, goods are tracked, supply chain actors are kept up to date, and all relevant information is shared in real time (Cui & Idota, 2018). At this point, it can be stated, that the outlined approaches can only cover a fraction of what has been studied in detail in previous research approaches. It is also not the goal of our research to develop a fully comprehensive overview of data-driven approaches in the procurement and supply chain

Table 1 Selection of previous research findings and classification of our approach

Exemplary research approaches: Data as a starting point for decision-making	
Hannila et al. (2022)	Approach of analyzing company data assets for improving data-driven decision-making
Kassner et al. (2015)	Product Life Cycle Analytics Approach with focus on structured and unstructured data
Phillips-Wren et al. (2015)	Big data analytics framework including the necessary process and related components
Zakir et al. (2015)	Big data analytics systems to predict volumes and actions for better decision-making
Zrenner et al. (2017)	Data source taxonomy for creating an end-to-end visibility of data
Exemplary research approaches: Relevant technical aspects of data for decision-making	
Bellatreche et al. (2022)	Data repositories and data models as an important success factor for data value
Garrido et al. (2022)	Privacy-enhancing technologies (PETs) and relating automotive use cases
Johanson et al. (2014)	Knowledge-driven product development with the help of vehicle data
Leitner et al. (2014)	Data mining for enhancing quality management in automotive manufacturing
Luckow et al. (2018)	Evaluation of deep learning techniques in automotive manufacturing domain
Exemplary research approaches: Data analytics in procurement and supply chains contexts	
Baryannis et al. (2019)	Approach of using machine learning for predicting supply chain risks on real case
Brintrup et al. (2020)	Approach of using data analytics for predicting supplier disruptions
Cui and Idota (2018)	Approach of reconfiguring supply chains with the help of blockchains
Ellram and Tate (2016)	Analysis of secondary data and its potentials for decision-making
Handfield et al. (2019)	Investigation of analytics landscape in the field of procurement
Vieira et al. (2019)	Big data approach for real-time supply chains by simulating scenarios
Classification of our research in the current research priorities	
Identifying relevant data sets and its usability for supporting decision-making in procurement departments by building on previous research approaches on classifying data and on data analytics in the fields of procurement or supply chains	

context. Rather, we want to use the examples to illustrate that data itself is the necessary prerequisite for any of these approaches to work. However, what initially sounds self-evident and almost self-explanatory is not common practice in organizations. In many organizations, projects for implementing data analytics fail because data is missing, or the quality of existing data is not sufficient to work with it (Handfield et al., 2019). Therefore, our research approach focuses on the relevant data in the procurement environment. With the help of expert interviews, we extracted relevant data that is either already being used in the respective organizations or that needs to be procured in order to improve decision-making. Many insights came from crises triggered by the COVID-19 outbreak. When supply chains suddenly broke down and decision-makers in procurement departments realized that massive deficits regarding necessary data had been uncovered. Thus, on the one hand, we use the already existing research approaches for the formulation of targeted questions to the experts and, on the other hand, we support these research approaches with our study by

uncovering relevant data and their usability. Table 1 presents existing research approaches and categorizes our approach.

3 Method

3.1 Research Framework and Setting

Based on our central research question, we use three input factors. First, previous research findings in the fields of data and data structures are relevant for a sound understanding of how to classify relevant data in procurement departments. Building on this, we used previous research in the area of data analytics for decision-making in procurement departments to review which approaches have already been developed by authors. The previous knowledge also served as a basis for the development of our interview study. The results of the interviews represent the central element of our research. For the analysis of the interview results, we followed the procedure of Gioia et al. (2013) which was used

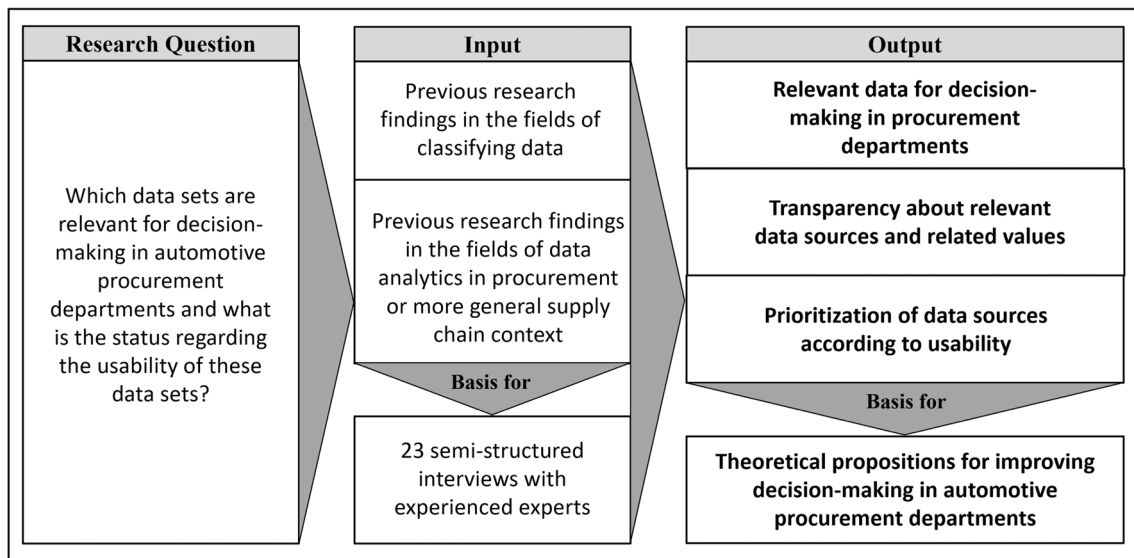


Fig. 1 Research framework

for the successive analysis and processing of our qualitative data. For this purpose, we first worked with an open coding scheme to extract and categorize all relevant data mentioned by the experts. The results will be explained in detail in the next section. These results serve as a basis to derive propositions for future research as well as practical application in organizations. Figure 1 demonstrates our research framework and the individual elements from the central research question to the planned output.

3.2 Data Collection

External shocks like the COVID-19 outbreak or the war in Ukraine disrupted numerous supply chains and many organizations had to make important decisions while struggling with massive uncertainty. At the outset of our paper, we explained the particular challenges faced by automotive manufacturers and the role of their procurement departments, especially impacts of external shocks and further crises (Bag et al., 2023; Gölgeci et al., 2023). Therefore, the data deficits brought to light by such external shocks and other crises were an important part of our expert interviews. Our interview questions addressed different areas of focus, which can be categorized as follows:

- Person, role and tasks in the organization.
- Relevant data in procurement departments.
- Data availability and data quality in procurement departments.
- Optimization of data availability and data quality in procurement departments.
- Relevant data competencies in procurement departments.

Above all, the relevant data in procurement departments and related availability or usability were of primary interest for this research. In this context, the experts also talked about the technical challenges at many points in the interviews. Our interviewees are experts in the fields of data analytics and business intelligence and also have experience in the procurement context or in a more general supply chain context. We used contacts from past or ongoing data analytics projects to conduct 23 interviews from 20 different companies. 45% of the companies have more than 100,000 employees, 25% have between 1,000 and 10,000, and 30% have less than 1,000 employees. The interviewed experts work for three different automotive manufacturers, for engineering and technology companies, for consulting companies or for software companies and data experts. We were therefore able to assemble a very broad field of expertise. Some of our experts have a technical background, others are in higher management levels of the consulting companies or executives in their companies, so that numerous facets of experience are covered as well. 35% of our experts have more than 10 years of professional experience, 74% more than 7 years. This ensures that the experts have profound experience that was valuable for the interviews. Table 2 presents the interview data in detail. Furthermore, we present information of our conducted interviews, for example interview durations or interview format. Thus, the interviews were mostly conducted online with the help of suitable software. In some cases, however, we were also able to conduct the interviews on site.

For the preparation of the interviews, we sent out an interview guide in advance. This guide contained the context of the interviews and the planned questions. These were asked

Table 2 Interviewee data

ID	Job Title/Position	Company	Employees	Experience (in Years)	Duration
1	Head of data platforms and solutions	IT consulting 1	< 5,000	> 15	29:24 ^a
2	Senior data analyst controlling	Automotive manufacturer 1	< 50,000	> 10	31:21 ^a
3	Data application manager	Automotive manufacturer 1	< 50,000	> 7	30:41 ^a
4	Senior data management architect	Automotive manufacturer 1	< 50,000	> 5	35:03 ^a
5	Senior expert data analytics	Automotive manufacturer 2	> 100,000	> 7	28:57 ^b
6	Head of data analytics procurement	Automotive manufacturer 2	> 100,000	> 7	39:24 ^b
7	Senior project manager analytics	Automotive manufacturer 3	< 100,000	> 5	44:14 ^b
8	Senior manager data strategy	Strategy consulting 1	< 25,000	> 7	45:12 ^a
9	Manager business intelligence	IT consulting 2	< 1,000	> 10	32:35 ^a
10	Partner data analytics	Strategy consulting 2	< 50,000	> 7	34:09 ^b
11	Senior data analyst procurement	Software corporation	> 100,000	> 10	32:04 ^a
12	Project manager data analytics	Engineering/technology 1	> 100,000	> 10	30:32 ^b
13	Manager data strategy	Auditing/strategy consulting 1	> 100,000	> 5	37:02 ^b
14	Team lead data strategy	Data platform supplier	< 1,000	> 3	28:33 ^b
15	Key account	Data software supplier 1	< 1,000	> 7	28:44 ^b
16	Senior manager analytics	Strategy consulting 3	> 100,000	> 7	38:56 ^b
17	Lead developer automotive	Data software supplier 2	< 1,000	> 5	29:11 ^b
18	Manager analytics	Strategy consulting 4	< 1,000	> 5	35:04 ^b
19	Senior manager business intelligence	Auditing/strategy consulting 2	> 100,000	> 7	29:38 ^b
20	Senior data scientist	Engineering/technology 2	> 100,000	> 7	34:55 ^b
21	Manager data strategy automotive	IT consulting 3	< 1,000	> 15	36:14 ^b
22	Lead data scientist	Engineering/technology 3	> 100,000	> 10	27:55 ^b
23	Associate partner business intelligence	Auditing/strategy consulting 3	> 100,000	> 10	36:17 ^b

^aInterview was conducted on site ^bInterview was conducted online

successively in the interviews, always in the same order. However, we did not interrupt any discussions that arose, so that all aspects of content could be included. All interviews were conducted in German, as German was the native language of all experts. All questions were open-ended and based on findings from previous research, as explained in Section 2. We used the findings according to Schultze and Avital (2011) for creating our questions. At the end, we gave the opportunity to mention any other points that were not part of the initial questions, so there was a possibility to share further aspects and experiences. This opportunity was used by some experts, which led to further insights. All interviews were recorded with the consent of the interviewees and subsequently transcribed for final analysis of the results.

4 Discussion of Research Findings

As already mentioned before, the conducted expert interviews were the central input for our research. We analyzed all the details mentioned by the experts and structured the data. We also used the method of Gioia et al. (2013) by developing 1st order concepts, then 2nd order concepts,

and finally aggregated dimensions that reflect the themes and phenomena mentioned in the interviews. To achieve theoretical and practical saturation according to Glaser and Strauss (1967), we worked with open coding what led to increasingly aggregated dimensions for all relevant items. As a result, we determine three aggregated dimensions on basis of the expert statements, which represent the basis for our remarks, while they address the data-relevant core problems of automotive procurement departments. Figure 2 represents our results by illustrating our procedure and the connections between the steps. In the following subsections, we describe in detail the results by explaining the specifics and providing insights into the findings from the expert interviews. At the end of this section, we present theoretical propositions for improving decision-making and the application of our results in automotive procurement departments.

4.1 Relevant Procurement Data for Decision-Making

The first aggregate dimension of our qualitative analysis concerns the relevant data in procurement departments. The interviewed experts emphasized the importance of relevant data for qualitative decision-making, while clarifying that

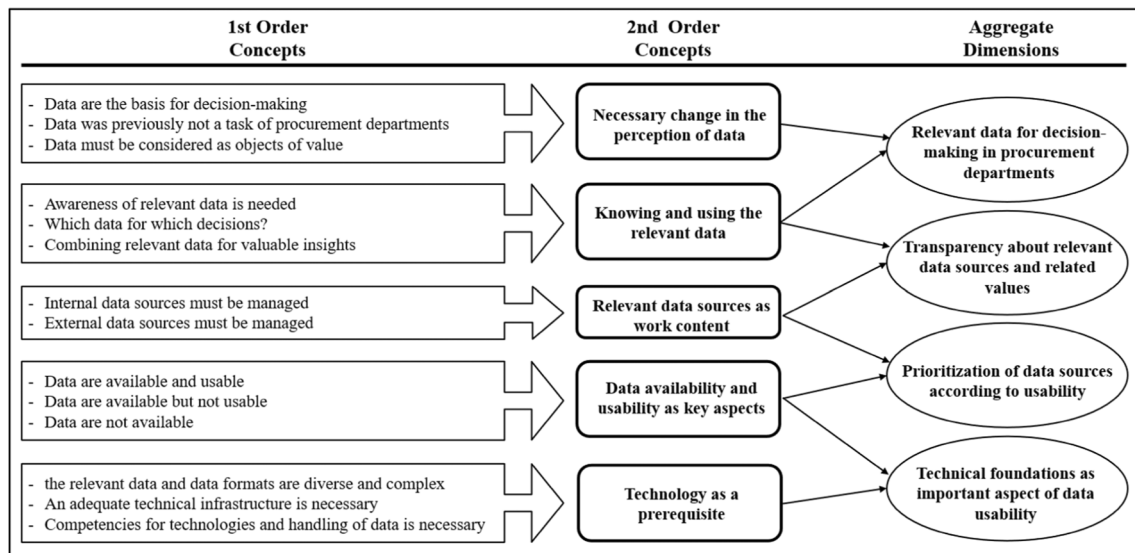


Fig. 2 Research findings according to methodology of Gioia et al. (2013)

the management of data has not been the focus in procurement departments in recent decades. Several times, we got the feedback that data needs to be considered much more as an object of value. So, there seems to be a clear awareness of the relevance of data. In this context, it was also emphasized several times that the COVID-19 outbreak or the war in Ukraine were accelerators for this perception. For example, supply chains broke down very suddenly, but there was a lack of transparency about procurement-relevant data at many points in the decision-making processes. Above all, transparency about the parts and their components was a major deficit:

“In the future, procurement departments will have to deal even more with data on components. Which components are installed in the final parts, where do they come from, should they remain the responsibility of the 1st tier suppliers or do they want to exert more influence themselves? In other words, technical data in order to be more capable of acting in the event of crises or sustainability issues.” (Interview ID2, Senior data analyst controlling, Automotive manufacturer 1).

Furthermore, internal supplier cost data was the most frequently mentioned relevant data from the experts' point of view. Numerous data sets were named, which must be consulted in procurement departments for decision-making processes. Delivery cost data, strategic cooperation data or tools and equipment data are only a selection. Above all, the non-transparent delivery stages were a frequently discussed topic in the interviews. In many cases, it was not possible to analyze which sub-components were affected by the supply disruptions, because the automotive

manufacturers could only evaluate who the direct suppliers of their parts were. Many further supply stages of 2nd tier or 3rd tier suppliers are not stored in the systems and, thus, supply chains broke off very suddenly. As can be seen in the quote, this also reveals a strategic component. Automotive manufacturers must decide for themselves how transparent they want to make their supply chains. The possibility of leaving the subsequent supply chain stages to the 1st tier suppliers is also mentioned, but this presupposes that their own actions in crisis situations depend on these suppliers. Other data requirements related to the use of tools and equipment by suppliers for independent aftermarket productions. Such tools and equipment often belong to automotive manufacturers and are made available to suppliers. According to the experts, it is often not clear what output quantities suppliers have for the legally mandatory independent aftermarkets. Tools and equipment are a large value item and in the responsibility of procurement departments. Other supplier data, such as sustainability or corporate social responsibility data, is also becoming increasingly important for procurement departments, as they are responsible for supply chains, according to the experts. They need to focus on the sustainability footprint as well as on aspects of working conditions in the production plants of their own suppliers.

In addition to internal supplier data, experts emphasize a focus on external data if procurement departments want to optimize their decision-making for future challenges, as the past external shocks demonstrated. In this context, completely new data sets, that have not yet been evaluated and analyzed in its entirety in many procurement departments, are essential for decision-making:

“Rather, completely new data will become important if a procurement department wants to become or remain effective. Thus, market data become important. For example, new competitors or new technologies as new reference line. Again and again, new innovative competitors enter the market. Often with completely new production processes. In the coming decades, many innovations in cars will also come from suppliers. Procurement departments must know such things, in the best case before all other competitors. It requires a permanent strategic market screening.” (Interview ID23, Associate partner business intelligence, Auditing/strategy consulting 3).

Thus, the importance of market data was emphasized in the interviews. Data on competitors or new technologies illustrate the focus on movements in the market, which in the best case, can be positive drivers and in the worst case can threaten the existence of the organizations. At this point, the relevance of suppliers for important innovations was also emphasized. In the automotive industry in particular, many innovations for competitive advantages are created hand in hand with suppliers, what further illustrates the need for such market data analyses. Many experts see procurement departments in the management function here, as they are the direct link to suppliers, as Pellengahr et al. (2016) point out. Certainly, such market analysis can also be conducted by other institutions in organizations. The experts name the close cooperation of procurement departments and thereby intensive relationships to suppliers. Thus, procurement departments can transfer innovations into the organization. Subsequent technical evaluations must be conducted together with development departments, however, procurement departments can be the input channel due to their role, as experts pointed out.

In addition to suppliers as central partners for procurement departments, the experts also named other players within supply chains who have great strategic value and must therefore be in focus when it comes to relevant data:

“Supply chain data is necessary. There are many more players in the supply chain than suppliers. Transporters, customs stations, intermediate warehouses, ports of all kinds, for example ship and air. Many stakeholders are involved, but experience has shown that not much data can be evaluated. Data on the locations of the goods or data on customs clearance for example. Core activities at this point: optimization of supply chains. If damage occurs, there are questions like: where did it happen, on the truck, or ship, in the interim storage facility? Not to forget borders. There are so many data bottlenecks within supply chains.” (Interview ID18, Manager analytics, Strategy consulting 4).

In this context, the experts were primarily concerned with the route taken by parts and components from various production facilities around the world to the automotive manufacturers and their just-in-time or just-in-sequence production operations. In very few cases is it possible to sustainably collect data such as exact locations and associated delays, damages or other issues and evaluate them for optimization purposes. The responsibility often lies with partners such as transport service providers or other stakeholders and, thus, also the sovereignty of the relevant data from a procurement perspective. As the number of bottlenecks increases, so do the risks. As previously discussed with regard to data on parts and components, the other stakeholders in the supply chains must also come into focus if procurement departments want to optimize their data situation for their decision-making. The topic of risks is also the segue to other relevant data for procurement departments, which the experts emphasized very frequently and illustrated with numerous examples:

“Then, of course, risk data [...]. Many OEMs are now in projects with service providers who can provide data. COVID-19, semiconductors and further crises have massively accelerated the need. However, there is still a lot to be done in the future, because the major added value will only come when the risk data can be directly incorporated into the company’s own systems. So, getting the information that supplier 1 is currently in the middle of a flood disaster still says nothing about which parts are affected.” (Interview ID21, Manager data strategy automotive, IT consulting 3).

Among other things, the experts named risks on suppliers, economic risk data, cyber security risks and compliance risks. In connection with this, a problem with externally purchased risk data has been described several times. There are numerous providers of software solutions for risk data that offer similar functionalities. Automotive manufacturers enter their own suppliers and the providers collect risk data on these suppliers and make it available in processed form. However, the information content from this is only low for automotive manufacturers, as no further and detailed analyses can be created from the pure information if this external data is not linked to the internal organizational data. A memorable example was given in the quote above. The information that a supplier is affected by a flood disaster and, thus, the supply chain is interrupted is important and good. The main added value would come from ad hoc information about which specific parts and components are affected and how they could possibly be replaced. To do this, however, the external risk data on suppliers would have to be linked to the internal system data, which is usually not the case because this step is highly complex. Two of our interviewees (interview ID15 and ID17) were experts from such service

providers. They confirmed the problem to us and referred to ongoing projects to establish such systemic connections. However, due to the very heterogeneous system architectures of their customers, there can never be a standard functionality for this.

For all the new perspectives that procurement departments need to adopt from a data perspective in order to optimize their decision-making, the experts also emphasized very clearly that the focused data in the past will continue to play an important role in the future:

“Today, we are very operationally positioned in procurement departments. We pay attention to important deadlines, budgets, performance targets and, rather with secondary focus, on issues of the future [...]. It becomes clear where the focus is and this focus will always be part of procurement departments and, therefore, will always remain important.” (Interview ID6, Head of data analytics procurement, Automotive manufacturer 2).

Spend data, demand data and governance data were mentioned as examples that have always been in the focus of procurement departments and will continue to be important. The experts confirmed very good data quality, as a great deal of expertise has been built up in procurement departments over the past decades. Thus, these data have always been part of the defined tasks of procurement departments in the past. Experts pointed to a continued need for focus, as the tasks of procurement departments will expand in the future due to many uncertainties, but previous tasks and duties will not disappear.

Cost data, which, not surprisingly, has been emphasized several times by experts, also fit into the context of data that have already been important. However, it was noticeable that many types of cost data were named that had not previously been the focus of attention. According to the experience of the experts, the nature of cost data is changing:

“Procurement departments need cost data. I do not mean the components costs. Procurement departments already have such data sets. I mean cost data for decision-making of supplier changes. Cost data for make-or-buy decisions. Data for the onboarding of new suppliers. From my experience, all procurement departments need work on such topics nearly every day. But again and again they start from the beginning and cannot benefit from existing data. Such data sets must be stored in a structured and evaluable form. For example, supplier changes costs including relocations, onboarding processes of new suppliers, development costs of components in-house instead of buying them from suppliers.” (Interview ID8, Senior manager data strategy, Strategy consulting 8).

Component costs were also mentioned. However, many experts named cost types that are often not available in an analyzable form, although they would be available. Costs for relocations or onboarding costs for new suppliers for example. In other words, cost data from activities that procurement departments perform again and again, but do not collect and use for future analyses. According to the experts, many important types of costs are not even recorded, but they would be essential for future decision-making. For example, it is often not possible to quantify which costs could be saved by eliminating variants or how offshoring and nearshoring affect sourcing decisions from a cost perspective.

Figure 3 summarizes our results. We analyzed all aspects from our conducted interviews and coded the inputs for categorizing them according to contents. We obtained six categories of relevant procurement data sets from the perspective of the interviewed experts. For a better evaluation, we added the frequency of mentions in square brackets. We can state that the own supplier data, as we called this category, dominates. However, this is immediately followed by external market data.

4.2 Relevant Data Sources for Procurement Decision-Making and Related Values

The experts told us about different categories of data and widely varying access options. We have reconciled the experts' practical experiences with the theoretical classifications of data from our 2 section and synthesized on this basis Fig. 4 as a result. We present the data sets explained in the previous section by sources and set a focus on the availability of each source.

According to the experts, relevant own supplier data and cost data sets come from both internal and external sources. Relevant market data, supply chain data or risk data are mainly obtained from external sources. Internal datasets are part of the respective organizations, while external datasets may be closed for free access (Zrenner et al., 2017). Supply chain data such as transport service providers, certification data, shipping and airport data or legislations data are time-consuming but available according to the experts. However, there are also relevant data sets that are not accessible to procurement departments, but which are urgently needed. For procurement departments, this means that external data in particular is a challenge when experts see a lot of relevant data from sources like suppliers or other stakeholders:

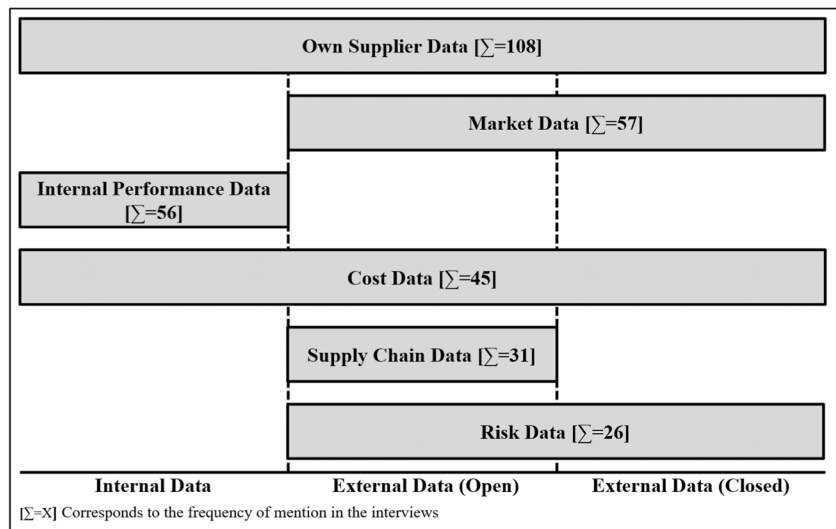
If you look at the major topics in procurement departments, it becomes clear that you will have to focus on strategic decision support in the future. Procurement departments need external data for this. In my opinion, the focus will be less on the purchasing performance

Own Supplier Data $\Sigma=108$	Delivery data [4] <ul style="list-style-type: none"> - Delivery stages data [9] <ul style="list-style-type: none"> - Sub-supplier data [8] - Production location data [7] - Delivery reliability data [14] - Missing, damaged, incorrect components data [3] - Terms of delivery data [2] - Delivery capacity data [2] - Incoterms data [1] Component data [1] <ul style="list-style-type: none"> - Component properties data [1] <ul style="list-style-type: none"> - Technical specifications data [2] - Warranty data [1] - Storability data [2] - Bill of materials data [6] - Component quality data [4] Sustainability data [3] <ul style="list-style-type: none"> - Sustainability footprint data [3] - Use of rare earth data [2] Corporate social responsibility data [2] <ul style="list-style-type: none"> - Internal compliance data [3] - Foreign working conditions data [2] Strategic cooperation data [1] <ul style="list-style-type: none"> - Contract data [3] - Onboarding data [2] - Performance measurement data [1] Invoice data [3] <ul style="list-style-type: none"> - Invoice totals data [2] - Terms of payment data [1] Tools and equipment data [1] <ul style="list-style-type: none"> - Location data [2] - Maintenance condition data [1] - Use for independent aftermarket data [1] Overall contract data [4] <ul style="list-style-type: none"> - Contact data [3] 	Supply Chain Data $\Sigma=31$ <ul style="list-style-type: none"> - Transport service provider data [13] - Country-specific certifications data [6] - Customs requirements data [5] - Shipping and airport data [3] - Interim storages data [2] - Legislation data [2] 	Cost Data $\Sigma=45$
	Risk Data $\Sigma=26$ <ul style="list-style-type: none"> - Supplier risk data[1] <ul style="list-style-type: none"> - Financial ratings data [6] - Management changes data [3] - Scandals and other incidents data [3] - Production downtime data [2] - Insolvency data [1] - Compliance data [1] - Sustainability risk data [1] - Economic crises data [2] - Political crises data [2] - Strikes data [1] - Cyber security risk data [1] - Internal compliance risk data [1] - Internal sustainability risk data [1] 	Internal Performance Data $\Sigma=56$	
	Market Data $\Sigma=57$ <ul style="list-style-type: none"> - Competitive data [3] <ul style="list-style-type: none"> - Emerging competitor data [10] - Product data [5] - Strategic alliances data [3] - Advancing other technology data [4] <ul style="list-style-type: none"> - Semiconductor technology data [4] - Storage technology data [1] <ul style="list-style-type: none"> - Cloud storage data [2] - Physical storage data [1] - Display technology data [3] - Necessary licenses data [3] - Advancing automotive technology data [6] <ul style="list-style-type: none"> - New sensors development data [2] - New battery development data [1] - New software development data [1] - Global supplier and service data [4] - New relevant patents data [3] - Global mergers and acquisitions data [1] 		

$\Sigma=X$ Corresponds to the frequency of mention in the interviews

Fig. 3 Relevant procurement data extracted and categorized from expert interviews

Fig. 4 Relevant procurement data and classification by data source



and more on making the supply chain transparent. To do this, you need data from suppliers, data from crises around the world. (Interview ID4, Senior data management architect, Automotive manufacturer 1)

We explained one important example in the previous section. The supply stages are often unknown to procurement departments. Sub-suppliers behind the 1st tier suppliers are often unknown for automotive manufacturers. At this point, the need for such supply chain transparency becomes obvious and, thus, appears as an important goal for procurement departments. According to the experts, cost data also

includes all levels of data sources. There is relevant data that is available internally, such as transport cost data, but there is also data that comes from external sources. Above all, component costs are provided by suppliers. These are initially externally closed. The level of detail to which they are made available to procurement departments depends on the type of collaboration between automotive manufacturers and their suppliers, as the experts emphasize. Market data and risk data show that there are many external closed data sets that confront procurement departments with major challenges. Competitor data or technological advances can often only be evaluated or used when they become public on the market,

and then they are very heterogeneous and very difficult to systematize, the experts stated. Risk data sets are a challenge mainly because it is too late when risks are public. Thus, the focus of risk data is on prediction. Many relevant market and risk data sets are external closed. Therefore, approaches are needed to open such sources. The experts emphasize the important role of procurement departments as managers of supply chains. The experts emphasize the important role of procurement departments as managers of supply chains. This role also involves the need to get important data in-house that is currently not available:

“As far as external data was concerned, for most of our customers, this was limited to supplier performance such as delivery times, delivery quality, and so on. COVID-19, semiconductor shortage, aluminum shortage, magnesium shortage or also storms, earthquakes and everything else that happened [...] in the last few years, however, show very clearly that procurement departments can no longer stay in their little nest, but must fly out to go on data hunts for better decisions. In other words, if procurement departments want to fulfill their role as managers of the supply chain in the future, they will have to bring many new external data sets in-house.” (Interview ID8, Senior manager data strategy, Strategy consulting 1).

What is described in the quote very figuratively as leaving one’s own nest and setting off on a data hunt contains a very serious message. It is about nothing less than a completely changed self-perception including completely new tasks of procurement departments if a better decision-making in future crises or other challenging situations is the goal. One possible approach to meeting this goal was already addressed by discussing the collaboration of automotive manufacturers with own suppliers. It is summarized very concretely in the following quote:

“Procurement departments must pay attention on data that is located at the suppliers. This can only be achieved through cooperation. With Catena X, for example, there is a major project in which a large part of the automotive industry wants to do just that, but of course there is still a long, long way to go.” (Interview ID2, Senior data analyst controlling, Automotive manufacturer 1).

The experts mentioned the need for cooperation several times. If relevant data for procurement departments comes from numerous stakeholders or other stages of supply chains, but is not always accessible, a data cycle can be enabled via cooperations, as the experts suggested in different discussions. Five experts even suggested a systematic data processing of such collaborations, including precise performance measurements and a data-based contract management. The

idea behind this is that access to external closed data cannot be forced. Rather, the focus should be on achieving an exchange of data. In other words, a cycle between the stakeholders in the supply chain. If procurement departments are important co-creators of these supply chains, they should have also an interest in such goals. The Catena-X¹ project mentioned in the quote, in which numerous automotive manufacturers, suppliers and other important players want to create an independent data network, is a current example of how such problems can be addressed with the help of cooperation.

4.3 Usability of Data as Major Criteria for Decision-Making

Having first focused on the relevant data sets and the availability of the associated data sources in the previous sections, we now focus on the usability of data sets on this point. This was an important aspect mentioned by all 23 experts in the qualitative analysis. The experts emphasized several serious problems in this context. We created three usability categories based on the expert interviews:

- Data available and usable
- Data available but not usable
- Data not available

These categories are based on the practical experiences of the interviewed experts. We related these categories to previously discussed data sources in order to classify the relevant data sets. For illustration purposes, we assigned the ten most frequently mentioned data sets, as Fig. 5 shows.

The interviewed experts told us about difficult situations in projects or daily situations, in which certain data were not available at all. In some cases, important aspects of the data were missing. At least, no added value was possible and several projects failed:

“Currently, a lot of data is not available. Unfortunately, the data quality is also poor at the customer end, but that doesn’t change the needs. In fact, the necessary effort is what keeps many companies away. They are afraid of the technical effort, the costs, the hassle. So, the question is always how to get the different types of data together and then also draw the added value from it.” (Interview ID1, Head of data platforms and solutions, IT consulting 1).

¹ Homepage of the Catena-X project including vision, goals and current activities of the project. <https://catena-x.net/en/> [28.11.2022].

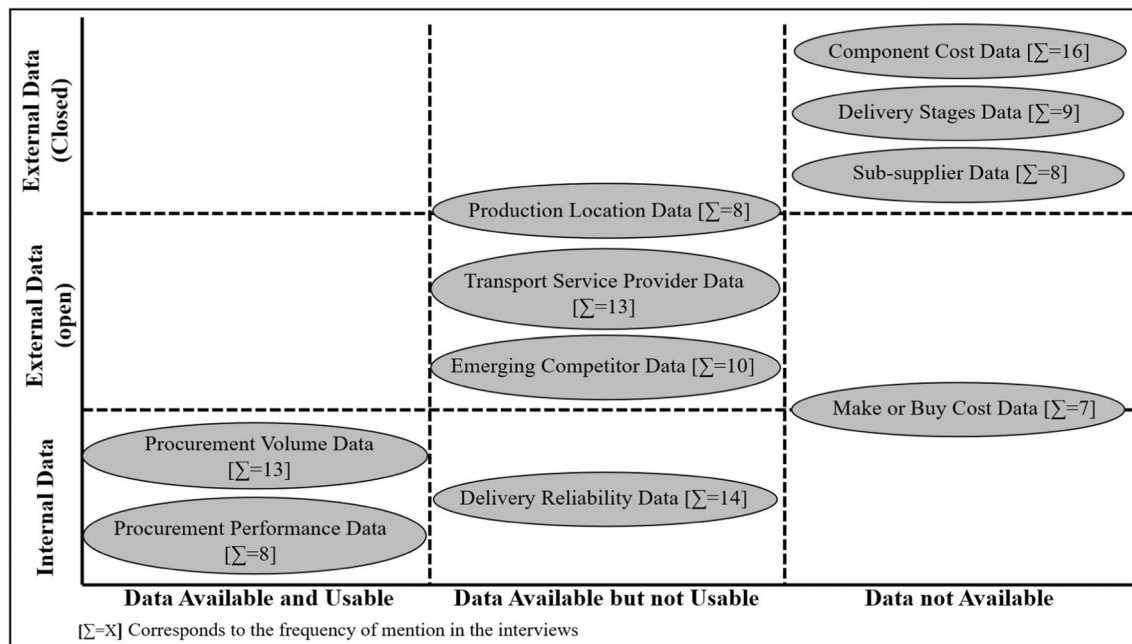


Fig. 5 Relevant procurement data categorized by source and usability

Often, the initial effort to generate complete and qualitative data is very high, but the associated added value is difficult to quantify as the experts told us. As a result, organizations invest in immediately measurable improvements, which is mostly accepted until certain requirements, for example through crises, make the massive deficits transparent, as the outbreak of COVID-19 did, when supply chains broke down and everyone was calling for reports and data analysis of suppliers, production plants or affected parts and components, that were impossible to conduct due to numerous data deficits. Examples of this are the component cost data, which are only available as they are quoted by suppliers. Other examples are the delivery stages and sub-supplier data sets, which we have also discussed, as these are only transparent to a limited extent for automotive manufacturers, since the contracts are usually concluded exclusively with the 1st tier suppliers and do not take the subsequent stages into account. At this point, the problem for procurement departments becomes very clear. Data that was not previously available but is urgently needed is one challenge. Another challenge is data that is already available but cannot be linked to generate new insights due to quality or further issues:

“We always think in categories in our team. We have a lot of data and can use it, then we have data, but it still fails to add value because the individual data cannot be linked and used. And then there is data, that we need but simply don’t have.” (Interview ID20, Senior data scientist, Engineering/technology 2).

At this point, the experts told us about wasted potential, since the data is already available. Thus, there is no effort for data sourcing activities. Nevertheless, these problems were mentioned very frequently by the experts. According to the experts’ experience, the reasons for this are a lack of knowledge in dealing with data, decentralized data sovereignty or a missing priority for such activities. Examples from the most frequently mentioned data sets are production location data, transport service provider data or emerging competitor data. Much of these data sets are saved at respective locations, but often cannot be used for more in-depth analyses. One expert gave us the example of data on global transport service providers, which is saved locally for sourcing decisions, but it is not possible to link these local data to concrete further data. For example, if an alternative service provider is necessary in crisis situations and information on international locations in connection with production plants is required. In other words, data is often available locally for specific purposes with limited characteristics but cannot be linked to other organizational data. These problems also exist with internal data sets:

“In addition to external data that we often need but not have for the future decisions, there are also internal data sets, that we actually cannot put into usable forms.” (Interview ID7, Senior project manager analytics, Automotive manufacturer 3).

For example, two experts told us about delivery reliability data that are available internally for goods receipts in order to carry out an evaluation for warranty or recourse

purposes. These data sets are carried out as documented reports so that they cannot be analyzed or linked to further data systematically. The examples are certainly very subjective and cannot be applied to all organizations. At this point, we emphasize the overarching issue of local data that exists in organizations but cannot be linked to other data. From the perspective of data analytics, this is where obstacles arise, so that potentially many valuable new insights remain untapped. Therefore, all examples shown in Fig. 5 can only serve as an illustration of the problem descriptions. Changeable for procurement departments are not the data sources, but the usability categories. After creating a transparency over relevant data sets, concrete activities can be developed, to improve the usability of specific or maybe of all relevant data sets. Our overall goal is to identify relevant data and analyze the usability for supporting procurement departments and their decision-making processes. From our perspective, this approach can help to provide transparency about important data and how to make it usable.

4.4 Technical Aspects as Necessary Foundation

In addition to the data-related challenges discussed so far, massive technical challenges also exist for procurement departments, as the experts conveyed to us very clearly in the interviews. In this context, a major challenge is the fact that there is not just one data format or generally one type of data, but diverse structures and associated massive complexity in the storage and use of the various data:

“A huge issue in all projects is the right handling of data. There is not one data format. We pull a lot of data from the Internet, for example data on risks with web crawling. But we also use existing databases from providers like Dun & Bradstreet or others, or we often collect data ourselves through interviews with partners or benchmark workshops. [...] In 99% of cases, customers have classic business warehouse in conjunction with their ERP software. Hadoop and similar technologies are foreign words there. I can't just put everything into the classic business warehouses and get something out of it. In all my foreign projects, I have been with exactly one customer who tried to counter this flood of data with a data lake. But the complexity in the logic layer was so massive that it didn't work either. So, the infrastructures in the OEMs are often not designed for the data analytics potentials.” (Interview ID18, Manager analytics, Strategy consulting 4).

Managing such diverse types and formats of data also requires the necessary technology. However, this has not yet been implemented everywhere by automotive manufacturers or often does not get beyond the status of pilot studies. As a result, many procurement managers lack the technical

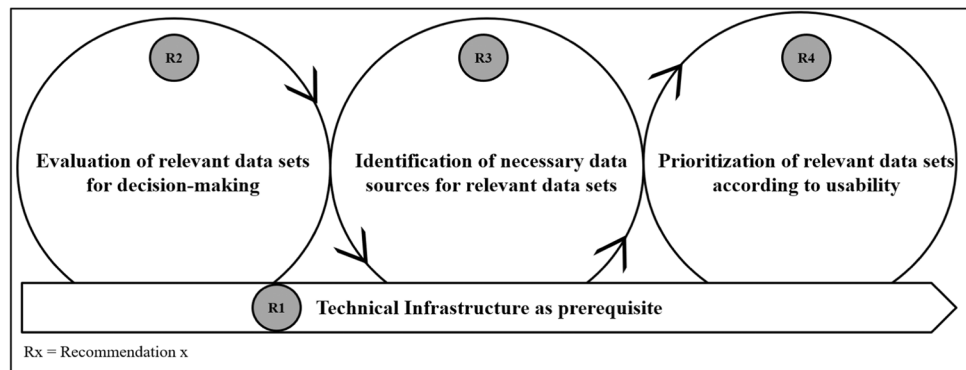
platform on which to perform all the relevant steps, such as analyzing the important data and making it usable:

“Automobile manufacturers usually do not have powerful data lakes or huge data warehouses. And what they do have is not suitable for storing everything. That is an issue. Big data and analytics live on infrastructure, and that requires different technology and also different competencies than before. Automotive manufacturers are not working intensively with data mining or in-memory solutions or Hadoop solutions or complex event processing. Here and there you can find pilot studies and very rarely real implementations are carried out. But often it fails because of the initial investments, which are indeed very high, but it is necessary to calculate the long-term potentials against it. These potentials are often difficult to quantify. And if you want to do analytics, but the basic structure is missing, you can imagine what that means from a data quality perspective.” (Interview ID21, Manager data strategy automotive, IT consulting 3).

Experts emphasized that other technologies than in the past must be considered, but that necessary competencies of responsible experts must also change as a result. However, experts from the OEMs also stated that the implementation of such technologies is the responsibility of IT departments and not procurement departments. According to these experts, the ability to assess such technologies and to formulate them as requirements in the direction of IT departments is the task of procurement experts. Therefore, the necessary competencies are also relevant for procurement departments. In this context, several experts mentioned very specific challenges with the data formats, which also strongly influence the selection and evaluation of necessary technology. Procurement departments need data from diverse sources and are therefore confronted with a wide variety of data formats. Thus, the experts also emphasized the difficulty of adhering to format standards due to the diverse data, but above all in handling the masses of data and the lack of knowledge within organizations:

“Not all data exported from the Internet or other external data, for example from sensors, can be compressed into a standard because of the various sources and data formats. In this context, it is also about data transfers. Masses of data must be transferred, and this requires know-how that is often not available. And for data protection reasons, it is also not possible to simply send all the data, for example the huge amount of sensor data, to existing and powerful applications, such as MongoDB, Apache Spark or Pentaho or something like that. We have not yet generated any knowledge and tried to build or own solutions, but these were not pow-

Fig. 6 Conducted process of identifying relevant procurement data and related recommendations



erful enough and failed in the pilot phase.” (Interview ID22, Lead data scientist, Engineering/technology 3).

In addition to the numerous technical challenges and related competencies already named, several experts told us about major challenges in dealing with data. For example, the competency within organizations of making relevant but very diverse data from numerous sources usable was named as a relevant challenge. At this point, the handling of diverse types of data and data formats is emphasized, but this can be seen as a prerequisite for dealing with data. In the following quotation, among other things, it is also recommended that the focus be placed more on data science and that data be made part of everyday life. The expert does not mean that procurement employees become data scientists, but rather that basic technical skills for dealing with data are built up:

“In my opinion, we need to move more in the direction of data science. I don’t mean that everyone has to be a data scientist, but organizations should think about how they can get people to deal with data. I’ve already said that it’s always about finding data, extracting it, preparing it and using it. In order to do that, I must be able to handle it. I must be able to link different types of data to generate added value. But organizations often don’t know Python or SQL, and there is no software like SPSS or whatever. And the data doesn’t fit into the existing data structures because there are classic ERP systems with the corresponding data sources. And, thus, massive data potentials are lost every day. It’s not enough to just think about data sources or just making them usable if there are no competencies for using diverse data formats appropriately.” (Interview ID19, Senior manager business intelligence, Auditing/strategy consulting 2).

Based on the expert interviews, various challenges for data-driven decision-making in procurement departments could be derived. From the identification of relevant procurement data to the utilization of this data and the necessary technical infrastructure, the experts spoke very openly

about current projects, challenges, and solution approaches. In the next section, we build on these insights and formulate recommendations for action for procurement leaders and their departments based on the findings.

4.5 Recommendations for Practical Application and Future Research

In this context, we developed recommendations for action based on our findings that are intended to emphasize the aspects of our work on the one hand and to serve as a guide and support for procurement departments on the other. Figure 6 summarizes our findings by illustrating our research process of identifying necessary technical aspects as a foundation, relevant data sets, related sources and the degrees of usability. Organizations are at different stages of development with regard to their infrastructure, the internal transparency of important data or the usability of this data. With our research, we aim to provide approaches based on the experience of experienced procurement and analytics experts. Depending on the stage of development, procurement leaders and their departments can therefore probably benefit to different degrees from our structured overviews and insights. We present practice-relevant recommendations for action, which in the best case can be applied in the everyday practice of procurement leaders, regardless of their stage of development. We discuss these four recommendations in detail in the following section.

Recommendation 1: Procurement departments rely on adequate technical infrastructure to enable data management for decision-making.

With the help of our expert interviews, we were able to highlight the importance of an adequate technical infrastructure as a necessary condition for procurement departments to be able to use data sources and the resulting data at all. This is less about specific generally applicable technology recommendations in detail and more about fundamental requirements for handling large volumes of data, diverse

data formats and varying data types. Each procurement department must presumably build up a technical basis individually. Such a finding is neither scientific news nor particularly surprising from the perspective of practitioners. During the expert interviews, it was found that this actual initial step has often not yet been taken. For example, the potential of data for decision-making in procurement departments is often emphasized and, in some cases, investments are made in software or specific reports. In many cases, however, the technical basis for being able to use the data is lacking. In this context it can also be stated that competencies are missing to suitable new technologies and handling data in procurement departments, although there is massive potential for organizations (Klee & Janson, 2022; Klee et al., 2021; Shao et al., 2022). At this point, it becomes clear that organizations and their procurement departments have to deal with their technical infrastructure and do the necessary preliminary work, so that subsequent steps such as the identification of relevant data sources and the resulting data or the usability of the data require this preliminary work.

Recommendation 2: Procurement departments must first evaluate their relevant data sets for creating the basis for data-driven decision-making.

With a view to the research on relevant data for different industries and organizations, our overview provides a possible approach for data structuring by using the example of automotive procurement departments. Based on the interviews, we developed a comprehensive data overview that reflects a structured insight into relevant data in the daily work of relevant experts and their tasks in six categories named supplier data, supply chain data, risk data, market data, cost data and performance data. Building on this, we promise new insights into data and structuring types if research is intensified here. From a practical point of view, we aim to help organizations that must identify relevant data at first. We were able to extract various data from the interviews. The focus was on automotive procurement departments. However, such overviews are probably also needed in other organizational departments or industries. Thus, our findings mainly support organizations. Especially those that want to improve their decision-making by facing various data deficits. Nonetheless, this is also an extension of previous research, as many findings serve to structure data. Therefore, we see it as necessary to further expand research on relevant data, as this can be considered the basis for all data-driven approaches. Without the relevant data, developments in the field of artificial intelligence, big data or data analytics in general are not possible. Therefore, we highlight the importance of identifying relevant data. What initially sounds self-explanatory is often not the case, especially in practice, as we learned in the expert interviews.

Recommendation 3: Procurement departments need to identify the relevant data sources, as external data represents a major challenge for organizations.

For research on relevant data sources, we promise valuable findings if these are more intensively combined with practical experience. This will result in validations of the approaches from a theoretical point of view and in usable tools for organizations for defining their relevant sources and related challenges for using them. As already explained in Section 2, data can come from internal, external open and external closed sources (Zrenner et al., 2017). With our results, we use the previously theoretical research findings and add a procurement perspective as a practical example and validation. Within the whole process, data sets are categorized into the respective characteristics after explanation by the experts. Transitioning to practice, the experts told us that a large part of the relevant data for automotive procurement departments is outside of own organizations. While internal performance data is unsurprisingly entirely internal and, thus, readily usable, sources for market data, supply chain data, or risk data must be obtained exclusively from external sources and often not available for procurement departments. Even with cost data, which at first glance appears to be the core work content of procurement departments, buyers must contend with data that is not readily available. While performance data such as purchasing performance can be collected internally, much cost data on components is in the hands of suppliers and, thus, part of negotiations. Market data on competitors or new technologies, for example, are external closed and only become usable when they are deliberately published. Risk data often only becomes usable when it appears in the media. Therefore, an intensified focus on necessary external data sources is an important success factor for improving decision-making. Procurement departments have historically been focused on performance measurement as experts told us in the interviews. A new focus is needed here. It can be deduced, that decision-making with a view to future challenges is only possible if this important data sets can be made available from external sources. This also reflects the complexity, as external data, above all external closed data, is a major challenge for procurement departments.

Recommendation 4: Procurement departments need to assess the usability of relevant data sets for prioritizing the potentials.

From a theoretical perspective, our results complement previous research approaches on data aggregation levels, data timeliness, or data sources and structures (Kassner et al., 2015; Phillips-Wren et al., 2015; Zakir et al., 2015; Zrenner et al., 2017). The level of data usability as

a prerequisite for decision-making sets a new perspective and was, to the best of our knowledge, not part of previous research findings in this context. Therefore, we propose valuable new insights by highlighting the factor of data usability for classifying, structuring and using data. From a practical perspective, we propose valuable guidance for organizations and their procurement departments as a result of analyzing data usability in depth. Experts told us, that relevant data for decision-making in procurement departments is often not available. Likewise, there is the case that data is available but not usable. In the previous section, we illustrated these cases using the ten most frequently mentioned data sets. Component cost data, delivery stages data, or sub-supplier data are external data that are closed and not readily available. Competitive data or data on transport service providers would be freely available but are often not usable in procurement departments because there is no way to collect this heterogeneous data and make it analyzable, for example with web crawling. If procurement departments place an increased focus on data that is available but not usable, rapid potentials for decision-making can be realized. Based on this, non-existing data can be focused subsequently. For example, by using the sources explained above. However, the effort required to obtain non-existent data is likely to be higher than making existing data usable as the experts emphasize. Therefore, we recommend the suggested sequence to realize potentials for decision-making as quickly as possible.

5 Contributions, Limitations & Conclusion

Our overarching research question described the goal to identify relevant data and its usability in procurement departments. In this context, we want to support procurement departments in decision-making. Especially in view of current or future challenges, managers need valuable data for qualitative decisions (Davenport, 2006). For this purpose, we first analyzed existing literature in the field of data analytics, technical aspects in this field and, building on this, further literature in the context of procurement and supply chain. The research results were also the basis for the content of our interview questions. We had the opportunity to interview 23 experts in the field of data analytics or business intelligence. All of them have a professional focus on procurement or the supply chain context. Thus, we were able to ensure relevant and up-to-date expertise. Based on a qualitative data analysis, we were able to extract numerous relevant data sets for procurement departments, which we categorized using open coding according to the research approach by Gioia et al. (2013). This was the basis for the subsequent categorizations according to data sources and

data usability, which we believe are highly relevant for procurement leaders and their departments to derive a prioritization approach.

In this way, we present relevant contributions from a practical perspective. Based on expert knowledge, we provided an overview of relevant data sets. In this context, we presented the relevant data sources and the degree of usability of these data sets. As an important prerequisite, we have also highlighted technical aspects based on expert knowledge and presented their importance for the use of relevant data. Thus, we were able to illustrate the associated challenges in exploiting data sources and discuss initial approaches to solving them with the approach of strategic partnerships. In the context of data usability, it also became clear from the experts' experiences that problems do not only exist with external data because of missing accessibility. There are also numerous challenges with internal data, which often cannot be linked with each other due to local characteristics and use. As a result, many potentials remain unused. We presented the entire process and added recommendations for action that we believe can be valuable to procurement leaders and their departments as they seek to gain transparency into relevant data and associated prioritization of use.

Even though the main focus of our paper was on the practical support, we were also able to make theoretical contributions. First, we used previous research findings as a baseline. We focused very specifically on procurement departments because of their central role in supply chains. Thus, we adopted a specific procurement perspective based on existing research findings. Building on this, we were able to derive characteristics of data usability based on the interviews, which enabled a new perspective on relevant data, especially in conjunction with the data sources. In addition to previous findings on possible data classifications based on data structures (Kassner et al., 2015), data sources (Zrenner et al., 2017), data movements (Hannila et al., 2022) or other characteristics, the degree of usability of data seems to be a major challenge in practice. Therefore, we proposed valuable contributions if such aspects receive attention in more intensive research.

At this point we also want to transparently present existing limitations of our research. First, it must be noted that research in data analytics and decision-making is very dynamic. Especially due to current global political or economic uncertainties, environmental disasters or pandemics, the calls for prediction or measurability are getting louder as research has recently shown the correlation between data management and competitive advantage (Chen et al., 2012; Huang et al., 2020; Olszak & Zurada, 2019; Popovič et al., 2018). As a result, research is intense and also fast-moving. This also affects our results. Data that we mark as relevant for procurement departments is also in constant flux. Thus, continuous research is required. This is especially relevant

for procurement leaders and their departments. A constant review of prioritized data for their own decision-making is necessary. Our presented data sets, the related data sources, data usability, and necessary technical prerequisites are a snapshot of the current situation. To complement these aspects, it must be noted that organizations vary in the extent to which they have designed their data-driven decision-making processes. Therefore, our results may also vary in the degree of usability. The interviewed 23 experts have their expertise primarily in the areas of procurement or rather higher-level supply chain. Furthermore, our findings originate from the automotive environment. Transferability to other procurement areas in other branches or even to other corporate divisions must be examined in further research. Finally, our study highlights the potential of data analytics implementation and aims to support this implementation by focusing on the real usability of data in procurement departments. Thus, we acknowledge that our research mainly focused on managerial aspects rather than practical implementation. Therefore, future research can build on this by developing specific data sets and repositories in the form of real or simulated data. By doing so, researchers can enhance the applicability of data analytics in procurement departments and provide further actionable and valuable insights for practitioners.

From our perspective, however, we were able to show which data is relevant in automotive procurement departments and what the degree of usability looks like. In this context, we were also able to highlight technical requirements based on expert knowledge and found out that many of these technical steps are still open in many organizations. With the practice-relevant recommendations for action we aim to propose a process of improving data-driven decision-making and in automotive procurement departments. Therefore, the recommendations for action describe a process from the identification of relevant technical necessities, the analysis of relevant data and data sources, and their practical use in the everyday work of automotive procurement departments. We emphasized at the beginning of this paper how uncertainty is becoming the norm (Ambulkar et al., 2015) and what that means, especially for organizations that depend on complex global supply chains. In this environment, procurement departments are the key interface (Pellengahr et al., 2016), which is why we believe research in this area is important and valuable.

Acknowledgements We express our gratitude to our interview partners who took a very important part in this study.

Author Contributions Both authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Sven Klee. The first draft of the manuscript was written by Sven Klee. Andreas Janson reviewed, commented and edited the first draft of the manuscript. Both authors reviewed and edited previous

versions of the manuscript. Both authors read and approved the final manuscript in its current state.

Funding Open Access funding enabled and organized by Projekt DEAL. The second author acknowledges funding from the Basic Research Fund from the University of St. Gallen (GFF).

Data Availability The datasets generated and analyzed during the current study are not publicly available due the fact that we assured confidentiality but are available from the corresponding author on reasonable request.

Declarations

Ethics Approval and Consent to Participate Verbal informed consent to participate in the interview study was obtained prior to the interview. The research project presented in this manuscript does not require an ethics approval.

Consent for Publication Verbal informed consent to publish results in an anonymized form was obtained prior to the interview.

Competing Interests The authors have no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahuja, T. S. A., Ngai, Y. (2019). Shifting the dial in procurement. McKinsey. <https://www.mckinsey.com/business-functions/operations/our-insights/shifting-the-dial-in-procurement>. Accessed 21 May 2022.
- Ambulkar, S., Blackhurst, J., & Grawe, S. (2015). Firm's resilience to supply chain disruptions: Scale development and empirical examination. *Journal of Operations Management*, 33–34, 111–122. <https://doi.org/10.1016/j.jom.2014.11.002>
- Bag, S., Sabbir Rahman, M., Choi, T. M., Srivastava, G., Kilbourn, P., & Pisa, N. (2023). How COVID-19 pandemic has shaped buyer-supplier relationships in engineering companies with ethical perception considerations: A multi-methodological study. *Journal of Business Research*, 158, 113598. <https://doi.org/10.1016/j.jbusres.2022.113598>
- Baryannis, G., Dani, S., & Antoniou, G. (2019). Predicting supply chain risks using machine learning: The trade-off between performance and interpretability. *Future Generation Computer Systems*, 101, 993–1004. <https://doi.org/10.1016/j.future.2019.07.059>
- Bellatreche, L., Ordonez, C., Mèry, D., Golfarelli, M., & Abdelwahed, E. H. (2022). The central role of data repositories and data models in Data Science and Advanced Analytics. *Future Generation*

- Computer Systems*, 129, 13–17. <https://doi.org/10.1016/j.future.2021.11.027>
- Bethaz, P., & Cerquitelli, T. (2021). Enhancing the friendliness of data analytics tasks: an automated methodology. https://ceur-ws.org/Vol-2841/DARLI-AP_15.pdf. Accessed 29 Nov 2022.
- Blazquez, D., & Domenech, J. (2018). Big Data sources and methods for social and economic analyses. *Technological Forecasting and Social Change*, 130, 99–113. <https://doi.org/10.1016/j.techfore.2017.07.027>
- Brintrup, A., Pak, J., Ratiney, D., Pearce, T., Wichmann, P., Woodall, P., et al. (2020). Supply chain data analytics for predicting supplier disruptions: A case study in complex asset manufacturing. *International Journal of Production Research*, 58, 3330–3341. <https://doi.org/10.1080/00207543.2019.1685705>
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 1165–1188.
- Chiusano, S., Cerquitelli, T., Wrembel, R., & Quercia, D. (2021). Breakthroughs on cross-cutting data management, data analytics, and applied data science. *Information Systems Frontiers*, 23, 1–7. <https://doi.org/10.1007/s10796-020-10091-8>
- Cui, Y., & Idota, H. (2018). Improving supply chain resilience with establishing a decentralized information sharing mechanism. *Proceedings of the 5th Multidisciplinary International Social Networks Conference on - MISNC '18*, 1–7. <https://doi.org/10.1145/3227696.3227723>
- Darmont, J., Novikov, B., Wrembel, R., & Bellatreche, L. (2022). Advances on data management and information systems. *Information Systems Frontiers*, 24, 1–10. <https://doi.org/10.1007/s10796-021-10235-4>
- Davenport, T. H. (2006). Competing on analytics. *Harvard Business Review*, 84, 98–107.
- Dickhaut, E., Janson, A., Söllner, M., Leimeister, J. M. (2023). Lawfulness by design – development and evaluation of lawful design patterns to consider legal requirements. *European Journal of Information Systems*, 1–28. <https://doi.org/10.1080/0960085X.2023.2174050>
- Duan, L., & Da Xu, L. (2021). Data analytics in industry 4.0: A Survey. *Information Systems Frontiers*, 1–17. <https://doi.org/10.1007/s10796-021-10190-0>
- Ellram, L. M., & Tate, W. L. (2016). The use of secondary data in purchasing and supply management (P/SM) research. *Journal of Purchasing and Supply Management*, 22, 250–254. <https://doi.org/10.1016/j.pursup.2016.08.005>
- Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research. *Organizational Research Methods*, 16, 15–31. <https://doi.org/10.1177/1094428112452151>
- Glaser, B. G., & Strauss, A. L. (1967). The discovery of grounded theory: Strategies for qualitative research. *Routledge*. 9780203793206
- Gölgeci, I., Gligor, D. M., Bayraktar, E., & Delen, D. (2023). Reimagining global value chains in the face of extreme events and contexts: Recent insights and future research opportunities. *Journal of Business Research*, 160, 113721. <https://doi.org/10.1016/j.jbusres.2023.113721>
- Hallikas, J., Immonen, M., & Brax, S. (2021). Digitalizing procurement: The impact of data analytics on supply chain performance. *Supply Chain Management: An International Journal*, 26, 629–646. <https://doi.org/10.1108/SCM-05-2020-0201>
- Han, L., Hou, H., Bi, Z. M., Yang, J., & Zheng, X. (2021). Functional requirements and supply chain digitalization in industry 4.0. *Information Systems Frontiers*. <https://doi.org/10.1007/s10796-021-10173-1>
- Handfield, R., Jeong, S., & Choi, T. (2019). Emerging procurement technology: Data analytics and cognitive analytics. *International Journal of Physical Distribution & Logistics Management*, 49, 972–1002. <https://doi.org/10.1108/IJPDLM-11-2017-0348>
- Hannila, H., Silvola, R., Harkonen, J., & Haapasalo, H. (2022). Data-driven begins with data; potential of Data assets. *Journal of Computer Information Systems*, 62, 29–38. <https://doi.org/10.1080/08874417.2019.1683782>
- Huang, C. K., Wang, T., & Huang, T. Y. (2020). Initial evidence on the impact of Big Data implementation on firm performance. *Information Systems Frontiers*, 22, 475–487. <https://doi.org/10.1007/s10796-018-9872-5>
- Johanson, M., Belenki, S., Jalminger, J., Fant, M., & Gjertz, M. (2014). Big automotive data: Leveraging large volumes of data for knowledge-driven product development. *IEEE International Conference on Big Data (Big Data)*, 2014, 736–741. <https://doi.org/10.1109/BigData.2014.7004298>
- Kassner, L., Gröger, C., Mitschang, B., & Westkämper, E. (2015). Product life cycle analytics – next generation data analytics on structured and unstructured data. *Procedia CIRP*, 33, 35–40. <https://doi.org/10.1016/j.procir.2015.06.008>
- Klee, S., Janson, A., Leimeister, J. M. (2023). Automotive manufacturers and their stumble from one supply crisis to another: Procurement departments could be the game changer by using data analytics, but... *Hawaii International Conference on System Sciences (HICSS)* 2023. <https://www.ssrn.com/abstract=4238820>
- Klee, S., Janson, A., Leimeister, J. M. (2021). How data analytics competencies can foster business value – A systematic review and way forward. *Information Systems Management*, 38(3), 200–217. <https://doi.org/10.1080/10580530.2021.1894515>
- Klee, S., Janson, A. (2022). Data analytics for effective decision-making in crises - identifying relevant data analytics competencies for automotive procurement departments. *ICIS 2022 Proceedings*.
- Leitner, T., Feilmayr, C., & Wöß, W. (2014). Optimizing reaction and processing times in automotive industry's quality management. *Data Warehousing and Knowledge Discovery DaWaK*, 8646, 266–273. https://doi.org/10.1007/978-3-319-10160-6_24
- Li, Q., & Liu, A. (2019). Big data driven supply chain management. *Procedia CIRP*, 81, 1089–1094. <https://doi.org/10.1016/j.procir.2019.03.258>
- Lind, L., Pirttilä, M., Viskari, S., Schupp, F., & Kärrä, T. (2012). Working capital management in the automotive industry: Financial value chain analysis. *Journal of Purchasing and Supply Management*, 18, 92–100. <https://doi.org/10.1016/j.pursup.2012.04.003>
- Luckow, A., Kennedy, K., Ziolkowski, M., Djerekarov, E., Cook, M., Duffy, E., et al. (2018). Artificial intelligence and deep learning applications for automotive manufacturing. *International Conference on Big Data (Big Data)*, Seattle, WA, USA, 3144–3152. <https://doi.org/10.1109/BigData.2018.8622357>
- Messina, D., Barros, A. C., Soares, A. L., & Matopoulos, A. (2020). An information management approach for supply chain disruption recovery. *The International Journal of Logistics Management*, 31, 489–519. <https://doi.org/10.1108/IJLM-11-2018-0294>
- Moretto, A., Ronchi, S., & Patrucco, A. S. (2017). Increasing the effectiveness of procurement decisions: The value of big data in the procurement process. *International Journal of RF Technologies*, 8, 79–103. <https://doi.org/10.3233/RFT-171670>
- Garrido, M. G., Schmidt, K., Harth-Kitzerow, C., Klepsch, J., Luckow, A., & Matthes, F. (2022). Exploring privacy-enhancing technologies in the automotive value chain. *Journal of Network and Computer Applications*, 207. <https://doi.org/10.48550/ARXIV.2209.05085>
- Olszak, C. M., & Zurada, J. (2019). Big data in capturing business value. *Information Systems Management*, 37, 240–254. <https://doi.org/10.1080/10580530.2020.1696551>
- Pellengahr, K., Schulte, A. T., Richard, J., & Berg, M. (2016). Einkauf 4.0 – Digitalisierung des Einkaufs. *Vorstudie des Fraunhofer IML und Bundesverband Materialwirtschaft, Einkauf und Logistik e.V. (BME)*.
- Phillips-Wren, G., Iyer, L. S., Kulkarni, U., & Ariyachandra, T. (2015). Business analytics in the context of big data: A roadmap

- for research. *Communications of the Association for Information Systems*. <https://doi.org/10.17705/ICAIS.03723>
- Popović, A., Hackney, R., Tassabehji, R., & Castelli, M. (2018). The impact of big data analytics on firms' high value business performance. *Information Systems Frontiers*, 20, 209–222. <https://doi.org/10.1007/s10796-016-9720-4>
- Redman, T. C. (2008). Data driven: Profiting from your most important business asset (General management). *Harvard Business School Press*. 9781422119129
- Reinzel, D., Gantz, J., & Rydning, J. (2018). The digitization of the world. From Edge to Core. <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>. Accessed 30 Jul 2022.
- Schultze, U., & Avital, M. (2011). Designing interviews to generate rich data for information systems research. *Information and Organization*, 21, 1–16. <https://doi.org/10.1016/j.infoandorg.2010.11.001>
- Sengupta, T., Narayanamurthy, G., Moser, R., Pereira, V., & Bhat-tacharjee, D. (2022). Disruptive technologies for achieving supply chain resilience in COVID-19 era: An implementation case study of satellite imagery and blockchain technologies in fish supply chain. *Information Systems Frontiers*, 24, 1107–1123. <https://doi.org/10.1007/s10796-021-10228-3>
- Shao, B. M., Louis, S., Corral, R. D., K., & Li, Z. (2022). Best practices for leveraging data analytics in procurement. *MIS Quarterly Executive*, 131–142. <https://doi.org/10.17705/2msqe.00062>
- Shi, Y., van Toorn, C., & McEwan, M. (2023). Exploration–Exploitation: How business analytics powers organisational ambidexterity for environmental sustainability. *Information Systems Journal*. <https://doi.org/10.1111/isj.12456>
- Someh, I., Wixom, B., Davern, M., & Shanks, G. (2023). Configuring relationships between analytics and business domain groups for knowledge integration. *Journal of the Association for Information Systems*, 24, 592–618. <https://doi.org/10.17705/1jais.00782>
- Vieira, A. A., Dias, L. M., Santos, M. Y., Pereira, G. A., & Oliveira, J. A. (2019). Real-time supply chain simulation: A big data-driven approach. *2019 Winter Simulation Conference (WSC)*, 548–559. <https://doi.org/10.1109/WSC40007.2019.9004717>
- Zakir, J., Seymour, T., & Berg, K. (2015). Big data analytics. *Issues in Information Systems*, 16, 81–90. https://doi.org/10.48009/2_iis_2015_81-90
- Zrenner, J., Hassan, A. P., Otto, B., & Gomez, M. J. (2017). Data source taxonomy for supply network structure visibility. *Digitalization in Supply Chain Management and Logistics: Smart and Digital Solutions for an Industry 4.0 Environment. Proceedings of the Hamburg International Conference of Logistics (HICL)*, 23, 117–137.

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

Sven Klee is a researcher and Industrial PhD student at the Information Systems department and Research Center for Information Systems Design (ITeG) at the University of Kassel, Germany. He obtained his M.A. in information systems from the University of Kassel. He works as a project manager for digitalization projects in the procurement department of a German car manufacturer. His research interests focus on data management and data competencies in organizations with a special interest in the field of procurement and relevant application cases.

Andreas Janson is a Postdoctoral Researcher at the Institute of Information Management (IWI-HSG) at the University of St. Gallen, Switzerland. He obtained his PhD from the University of Kassel, Germany. His research focuses on service design, smart personal assistants, decision-making in digital environments, and digital learning. His research has been published in leading information systems and management journals such as the *Journal of the Association for Information Systems*, *European Journal of Information Systems*, *Journal of Information Technology*, and *Academy of Management Learning & Education*.