In order to understand how organizations may effectively leverage data as strategic resources, we need to consider how the data shape strategic processes and outcomes. We study the process of data-driven strategizing at a European postal service organization and focus on identifying the mechanisms through which data can influence strategic choices and actions. Our findings reflect two tensions regarding the role of data in data-driven strategizing. First, our findings affirm that while data may allow for strategic exploration by opening up opportunities beyond their historical purpose, they may also cause organizations to become trapped into staying close to the historical purpose for which they have been collected. We highlight that past choices and actions can become inscribed in the data, and that this historically situated nature of data can influence current and future choices and actions. Second, we find that, because of the nature of the data, organizations may be encouraged to closely collaborate with external stakeholders, while the data can also prevent stakeholders from becoming fully engaged. We highlight the need for data to be “contextualized” such that they can become meaningful to different stakeholders. Altogether, our findings cast criticism on a number of assumptions underlying the literature on data-driven value realization.

This paper has been co-authored by M. H. Rezazade Mehrizi; M. Huysman, and F. Feldberg.
3.1 Introduction

Anecdotal evidence and high-profile examples around “big data” (e.g., Marr, 2016) have encouraged managers to try and see how they can use data to benefit their organizations. Consequently, organizations have been acquiring data and data-related technologies, hoping that this will help them improve their performance (LaValle et al., 2011), increase their profits and market share (McAfee and Brynjolfsson 2012; Tyagi 2003), strengthen their competitive positions, and gain a competitive advantage (Davenport, 2006). Generally, data are considered strategic resources that organizations can exploit, not only to support their existing business strategies, but also to develop whole new, data-driven strategies for which data are leveraged “to create differential value” (Bharadwaj et al., 2013, p. 472).

Scholars too have been triggered by the potentials of data as strategic resources, and have examined how organizations may effectively exploit data in the development and implementation of data-driven strategies. Hereby, studies have largely discussed what resources and capabilities need to be in place. For example, studies have explored what challenges organizations may encounter in implementing data-driven strategies (Sivarajah et al., 2017), suggested multiple success factors to ensure that value can be realized from data (Chen et al., 2015; Davenport, 2006; Gao et al., 2015), and identified various technological and organizational capabilities that organizations need to strive for when trying to exploit data (Tallon et al., 2013-14; Wang & Hajli, 2017; Zeng & Glaister, 2017).

It seems that thus far, studies on this topic have mostly paid attention to the traits of organizations, the need for sophisticated systems, and the roles of human and organizational actors in processes of data-driven strategy development and implementation. By doing so, however, the data themselves have naturally been pushed to the background, and have been treated as if they can always be exploited to support any strategy. Consequently, we have limited knowledge of how the data may also influence strategic choices and actions as actors work with them in practice.

Some scholars have recently suggested that rather than pushing data to the background, we need to examine how data can shape the very processes aimed at exploiting them. For example, Tempini (2017) argues that “to understand the social implications of big data and datafication, we need to put data and their processes at the center of our analyses” (p. 192). Similarly, Jones (2018) has criticized existing conceptualizations of data as indistinguishable and decontextualized resources. He argues that we need to differentiate between different types of data and aim to understand how these data actually inform choices and actions. The arguments provided by these scholars are very
much in line with a stream of literature that has called for incorporating the role of technological and digital artifacts in organizing processes (e.g., Leonardi, 2010; Orlikowski, 2007; Vaara & Whittington, 2012). It seems that data as digital resources can have distinct characteristics that enable or constrain strategic choices and actions, thereby shaping strategic processes and outcomes (Leonardi, 2010; Vaara & Whittington, 2012). Following this reasoning, we argue that by bringing data to the forefront and unpacking how data influence strategic choices and actions, we can gain a better understanding of how organizations may effectively leverage data as strategic resources.

To help us understand how data shape strategic processes and outcomes, we adopt a strategizing perspective (Burgelman et al., 2007; Galliers, 2007; Mintzberg & Waters, 1985). This means that we perceive the development and implementation of data-driven strategy as an emergent process, in which data and strategic choices and actions continuously shape each other over time. We contribute to the literature by focusing on the role of data and ask: “How do data shape the process of data-driven strategizing?”. To answer this question, we study the process of data-driven strategizing within a European postal service organization (LogiCo) and focus on identifying the mechanisms through which data influence strategic choices and actions. Based on the findings from our case study, we explicate two tensions concerning the role of data in data-driven strategizing. First, we find that while data may allow for strategic exploration by opening up opportunities beyond their historical purpose, they may also cause organizations to become trapped into staying close to the historical purpose. Second, we find that, because of the nature of data, organizations may be encouraged to closely collaborate with external stakeholders, while the data can also prevent stakeholders from becoming fully engaged. We contribute by showing that contrary to much of the literature on data and “big data”, data cannot always readily be transferred to different contexts, nor can they infinitely be explored to be used for many different purposes. We also contribute by highlighting the need for adopting a process view in order to be able to understand how organizations may effectively leverage data as strategic resources.

The paper is structured as follows. In the theoretical background, we first conceptualize how adopting a strategizing lens helps to capture the mechanisms through which data shape strategic processes and outcomes. In our methodology section, we report how we captured the process of data-driven strategizing within a European postal service organization (LogiCo). In the findings section, we describe four “strategizing moves” and explicate the mechanisms through which data shaped these moves in our case. Finally, we reflect on our findings and discuss a number of implications for theory and practice.
3.2 Theoretical Background

Several studies have proposed that when organizations aim to leverage data strategically, they first need to define a number of well-chosen goals, establish a data-driven vision, ensure that these goals and vision represent the business goals, and devise steps for attaining them (Akter et al., 2016; Gao et al., 2015; DalleMulle & Davenport, 2017; Seddon et al., 2017). In other words, they suggest the formulation of a clear data-driven strategy, that is, an “organizational strategy formulated and executed by leveraging [data] to create differential value” (Bharadwaj et al., 2013, p. 472). Such a data-driven strategy is not merely concerned about how to manage data at the functional level, but especially about how data analytics can be applied to produce business value (DalleMulle & Davenport, 2017; Seddon et al., 2017).

Granted, without some kind of data-driven strategy, data may sit idle and insights might never lead to actions (Lavalle et al., 2011; Shollo & Galliers, 2015). But by emphasizing the need to develop initial strategic plans, studies are implicitly assuming that organizational actors will have a clear idea of what the data look like in terms of different characteristics, and how these data will enable strategies in their specific context—or that it does not really matter as long as you have “big” data that can be used for many different purposes. Moreover, an implicit assumption is that these data and their role will also remain stable over time as organizations actually develop and implement data-driven strategies. In practice, this may not be the case.

If we aim to understand how data shape strategic processes and outcomes, we need to acknowledge that data are, in practice, dissimilar, contextual, and dynamic resources. First, data are dissimilar in the sense that the data used by one group of actors can substantially differ from the data used by other groups of actors (Jones, 2018; Kitchin & McArdle, 2016). For example, data may vary in content (e.g., what phenomena they are supposed to describe), in how they have been recorded and structured (e.g., at what level of granularity or detail), and in their associations with other data (e.g., whether they are accompanied by metadata). Scholars have argued that rather than assuming that all data are equal, we need to be sensitive to the ways in which data with different characteristics may enable or constrain different choices and actions in the same context (Jones, 2018; Kitchin & McArdle, 2016). Thus, to understand how data shape strategic processes and outcomes, we need to be specific about the characteristics of the data that actors use in a particular case, and be open to see how different (combinations of) characteristics may contribute to influencing strategic choices and actions.

Second, data are contextual, which means that the role of data can only be evaluated relative to “the context in which data is produced and used”
(Strong et al., 1997, p. 104). This point highlights that *the same* data can afford and constrain different choices and actions depending on the context in which they are used (Jones, 2018; Zeng & Glaister, 2017; Strong et al., 1997; Wang & Strong, 1996). For example, how “interpretable”\(^\text{15}\) data are and how this affects choices and actions highly depends on who is interpreting the data and for what purpose. To understand how data may shape strategic processes and outcomes, we thus need to examine what organizations aim to use the data for, who are actually working with the data (e.g., in terms of roles and skills), how actors work with the data to fulfil certain goals, and what other factors and contextual conditions (e.g., rules and regulation) may be of importance (Strong et al., 1997; Jones, 2018). By treating data as contextual resources, we may be able to see how the role of data differs under differing circumstances.

Finally, we need to consider that data, and the role of data, may also change over time as actors work with them in practice. For one, data are *dynamic* resources that typically do not remain constant over time; data can be extended, combined, deleted, abstracted, and continuously updated as actors work with them in practice and through other digital artifacts (Kallinikos et al., 2013; Yoo et al., 2010). Data as digital resources are “unstable and unbounded, resisting reification by their very nature” (Ekbia, 2009, p. 2558). To understand how data shape strategic processes and outcomes, we need to follow such changes in the data. Additionally, we should acknowledge that even when the data themselves do not change over time, actors may still continuously uncover new functions of the data as they interact with them (Aaltonen & Tempini, 2014; Constantiou & Kallinikos, 2015; Kallinikos et al., 2013; Zeng & Glaister, 2017). Adopting a process perspective (Pettigrew, 1990; Pettigrew, 1992; Langley, 2007) seems particularly useful for examining the changing nature and role of data over time.

To understand how data shape strategic processes and outcomes, we adopt a strategizing perspective. According to this perspective, organizations do not simply develop a strategic plan and then implement it; rather, strategy is emergent and continuously shaped in a process of strategizing (Galliers, 2007; Marabelli & Galliers, 2017; Mintzberg & Waters, 1985). Scholars adopting this perspective aim to capture both processes of strategy development and implementation, and the dynamic, changing nature of strategy content (Burgelman et al., 2017). Importantly, strategizing scholars acknowledge that technologies do not necessarily follow strategy; rather, technologies and strategic choices and actions co-evolve in a process in which technologies and digital resources can also influence strategic choices and actions (e.g., Dameron et al., 2015; Even more “technical” characteristics like volume may only be evaluated relative to the context in which the data are used; what is considered as high-volume data to a postal service organization may be nothing to an organization like Google.)
Huang et al., 2014; Karpovsky & Galliers, 2015). Thus, adopting a strategizing perspective seems especially useful for studying how data shape strategic processes and outcomes.

### 3.3 Methods

We adopted a single case-study design (Eisenhardt, 1989, Yin, 1981) and studied a rich case, i.e., that of an established organization that invested in data-related capabilities and aimed to develop data-driven strategies that would allow them to attain economic benefits.

#### 3.3.1 Empirical Setting

LogiCo is a European postal service organization\(^{16}\), traditionally active in the business-to-business postal market\(^{17}\). The organization focuses on the delivery of *addressed* mail, that is, mail items that have been tagged with an address label, including transactional mail (e.g., invoices), direct mail (e.g., advertisements), and periodic mail (e.g., magazines). LogiCo delivers these mail items to households on behalf of business clients such as banks and retailers.

Near the end 2011, LogiCo was one of few serious players in the business-to-business postal market. However, LogiCo’s shareholders and its board of directors realized that the market of physical mail items was shrinking; business clients were increasingly using other, digital channels to reach their customers. Therefore, the board commanded the development of a new strategy that should help LogiCo survive in light of this shrinking market. The CIO and CCO mainly took it upon themselves to design the new strategy, guided by an external strategy consulting company. The team created a forecast of the shrinking market and constructed worst-case, base-case, and best-case scenarios for its decline. Their aim was to predict how much volume would remain in 2017 for each type of mail that LogiCo was delivering, that is, transactional mail, direct mail, and periodic mail. The strategy team had estimated that while transactional and periodic mail might disappear, direct mail—i.e., sending advertisements that are tagged with an address label to specific households—would remain to play a relevant role in their business clients’ marketing mixes.

The strategy team imagined that they could use data to help their business clients find new customers, and that this would increase LogiCo’s market share in the direct-mailing part of the market. They were convinced that data could help them gain insights about the interests of people living on addresses. Using such insights, LogiCo could then help business clients find the addresses

\(^{16}\) This organization offered the funding for our research project on the influence of big data on business model innovation. We have been careful to remain objective and unbiased, and highlight that there has been no conflict of interest in this study.

\(^{17}\) Although they deliver mail items to consumers, they do so on behalf of business clients.
of prospective customers to whom they could send their advertisements. By offering such targeted advertising as a product, strategy team members hoped to increase LogiCo’s mail volume (because business clients would consider physical advertising as an effective channel and place more orders), and increase the chance that advertising resulted in purchases and new customers for business clients (measured by the conversion rate). This general idea became the starting point for LogiCo’s data-driven journey. In the findings, we will provide a narrative of what followed: how actors began to experiment with data, developed data-driven products and services, and moved into different strategic directions.

In describing our narrative, we will often aggregate actors to a higher level of abstraction. When we refer to strategy team members, we mean those at the executive level, strategy consultants, and the board of directors and shareholders that approve and can influence strategic planning. When we refer to development team members, we mean actors charged with the design and implementation of data-driven products, including members from the departments of business intelligence, new business, marketing, and sales. Sales team members include sales managers, (key) account managers, account executives, and product managers. Finally, analysts are the actors who collect and process data, and do the actual analyses.

3.3.2 Data Collection and Analysis

We followed an inductive, iterative process to collect and analyze data, in which we can distinguish two phases (see Figure 3.1). In the first round of data collection (September 2014-November 2015), our goal was to understand LogiCo’s business context and the way the organization had (historically) operated. To this end, we interviewed middle managers and executive team members, and asked them specific questions about their individual roles (e.g., their tasks and responsibilities, and whom they collaborate with); their departments and divisions (e.g., the structure of the department, technologies in use, rules and regulations in place, and running projects); LogiCo’s strategic position (e.g., competitors, overall mission, and key partners), and how LogiCo had been creating and appropriating value over time (e.g., products and services offered, revenue model and cost structure)—all interviews were recorded and transcribed18. Moreover, we recorded and made notes of meetings on a multitude of topics. We also collected public documents from the organization’s website (e.g., press releases, news items, and leaflets); obtained documents from the organization’s knowledge management system (e.g., strategy documents, organograms, and internal newsletters), and received documents from the interviewees. Finally, the first author spent one week inside the organization to get

18 With the help of research assistants.
a feeling of the organizational culture and atmosphere, and familiarized herself with the organization’s processes by spending one day sorting and transporting mail items.

We analyzed the data collected in the first phase by constructing a comprehensive description of the case: a single document in which we integrated the different data elements (Yin, 1981). We carefully examined each interview transcript and ordered quotes under general themes, such as organizational structure; organizational culture; strategic positioning; buyer and supplier relations; rules and regulations; finished and running projects, and challenges and controversies that interviewees faced. We went through our meeting notes and documentation to cross-check the findings so far, and added additional insights to the description. The resulting “case report” functioned as a baseline for the second phase and subsequent analyses.

In the second round of data collection, we paid closer attention to LogiCo’s data-driven journey and focused on collecting and making sense of longitudinal data (Pettigrew, 1990; Pettigrew, 1992; Langley, 2007). We identified and interviewed actors that had been involved in the development of LogiCo’s initial strategy in 2012; in shaping LogiCo’s strategy since 2012, and in the initiation, development, and selling of data-driven products. We asked interviewees specific questions about their role; how data-driven products and initiatives had come to be and developed over time; what goals different actors had in mind; what controversies and challenges they faced, and how data-driven initiatives did or did not result in value, including what kind of value, and for whom. Interviews were semi-structured as we used an interview guide, but allowed interviewees to engage in conversation and also asked probing questions to collect more data (Patton, 2002)—these interviews were also recorded and transcribed. As we conducted the interviews between the end of 2016 and the end of 2017, interviewees mostly reflected on their experiences retrospectively. Documentation therefore played a more prominent role in this phase, as we could use documents to validate claims and make sense of the various events. Next to collecting more general documentation, we received strategic and project documentation about data-driven projects and initiatives from interviewees. We continued to attend, record, and make notes of exploratory and more focused meetings, and later also attended “innovation sessions” in which participants spent one full day discussing innovative strategic directions for LogiCo.

We focused our analysis in the second phase on constructing a “case

19 Except for one, because the interviewee was concerned about the sensitivity of what he had said. We recorded the interview and made detailed notes.
20 Documents often consisted of sets of PowerPoint slides. We treat different versions of the same document as separate, because the nuances might provide useful information.
21 These are not recorded, but we made detailed notes.
"story" that specifically focused on depicting the story of data-driven strategizing at LogiCo. Comparable to what we did in the first phase, we carefully analyzed each transcript and integrated different data elements around themes (Yin, 1981). At this point, we focused on themes that were related to the organization's data-driven journey (e.g., designing a new strategy, acquiring capabilities, developing products and services, and revising the strategy) and put effort into chronologically ordering the themes. Meeting notes and documentation helped to specify some ambiguities around the context and meaning of ideas discussed in the interviews. The resulting “case story” reflects how LogiCo tried to create and appropriate value from data over time; who was involved and how; what the data looked like; what external and internal contextual factors and events happened, and what the (intermediate) outcomes of data-driven strategizing were. We then coded this document for: 1) data objects and their characteristics; 2) how different actors worked with the data (we call this data-related practices); 3) opportunities and challenges that actors faced, and 4) other potentially relevant contextual conditions, at different points in time.

We structured our codes around strategizing moves\textsuperscript{22}, which we define as periods in time in which actors draw on a certain set of data-related practices as they engage in data-driven strategizing. Data-related practices represent ways of working with data, such as aggregating, combining, and analyzing data (Hartmann et al., 2016; Tempini, 2017). In a way, strategizing moves represent periods of continuity (Pettigrew, 1990), reflecting the persuasion of a single strategy for a period of time. We critically examined, for each strategizing move, how different characteristics of data contributed to raising challenges and opportunities for actors involved; how this changed the way actors worked with data, and how this affected the data-driven strategy processes and outcomes.

\textsuperscript{22} We refer to strategizing moves rather than episodes (Hendry & Seidl, 2003) to emphasize our intention to capture meso-level organizational moves rather than micro-level actions (Jarzabkowski & Spee, 2009).
3.4 An Empirical Narrative of Data-Driven Strategizing

We structured our narrative around four strategizing moves (i.e., periods in time in which actors draw on a certain set of data-related practices): 1) experimenting with mail order data to offer targeted advertising; 2) selling external data and providing analytical insights; 3) offering profiles and a tool to engage business clients, and 4) connecting business clients and a data broker for targeted advertising\(^\text{23}\). An overview of the moves and the mechanisms through which data influenced these moves is depicted in Figure 3.2. In the next sections, we summarize each strategizing move and critically reflect on how data influenced strategic choices and actions in each move.

\(^{23}\) We stop our analysis after the beginning of the fourth move. This provided a natural stopping point as by then, LogiCo had decided not to do anything with data in-house anymore. Rather, they fully relied on partnerships with a data broker.
Figure 3.2 Summary of strategizing moves and the role of data.
### 3.4.1 First Strategizing Move: Experimenting with Mail Order Data to Offer Targeted Advertising

In Table 3.1, we summarize the first move, which data-related practices were associated with this move, what data the different actors relied on in this move, and what key challenges the actors faced.

**Move 1: Experimenting with Mail Order Data to Offer Targeted Advertising**

**Description:**
LogiCo might use mail order data to gain insights that are relevant for offering targeted advertising to marketers and business clients who aim to acquire new customers. The strategy team develops a strategy to formalize this idea and an analyst begins to experiment with mail order data to see if these data can indeed be useful.

**Data-Related Practices:**
- Analysts combine mail order data from different clients based on the address field.
- Analysts deduce personal insights about someone from the type of mail that person has received in the past months.

**Data:** Mail Order Data
Mail order data contain specific information about mail orders, such as the type of mail, who sends it, and who received it. These data are uploaded by business clients. Analysts can combine mail order data from different clients because these data elements all share an address field. Mail order data are stored for several months and databases are updated each time a new order is placed.

**Key Challenges:**
LogiCo faces data ownership issues and is concerned about the ethical implications of using mail order data to gain insights that are relevant for offering targeted advertising.

**Table 3.1 Summary of the first move.**

Triggered by the notion that the postal market was shrinking in 2012, LogiCo’s strategy team set out to find ways in which they might survive. The strategy team anticipated that “mail order data”, which the organization had anyway been collecting for the purpose of efficiently delivering mail items, might also reveal insights about households that could be relevant for offering targeted advertising. These mail order data contained information about the mail items (e.g., sender, receivers, type of mail) that LogiCo had sent to addresses on behalf of business clients. To support the efficient delivery of mail, business clients had to specify, for each delivery order: 1) what type of mail they wished to send (business clients would give a description), and 2) where they wished to have these mail items delivered (business clients would submit a list of addresses). Analysts could combine these data from different clients into one dataset specifying, for each address, what mail items had been delivered there. These data were saved for a limited period of months and were updated whenever a business client placed a new order.

In 2012, by means of an experiment, one of LogiCo’s analysts sought how much information could be gained about one of LogiCo’s executive team members, based on data about the mail items that this executive had received in prior months. The analyst grouped mail orders from similar clients (e.g., mail from insurance companies, mail from banks) and analyzed the descriptions of mail items that had been sent to the executive’s address. It turned out that

24 Many things can be said about the data. In the tables and the stories, we focus on those characteristics that are key to explaining how data shaped the process of data-driven strategizing.
the analyst could extract quite relevant information from these data, including information about the executive’s hobbies (e.g., if they received fishing magazines), the age of his children (e.g., from mail items sent by municipalities or insurance companies), and the type of car he drove (e.g., when he received a BMW magazine).

However, parallel to discovering what insights could be gained from mail order data, a legal expert explored whether LogiCo was, from a legal perspective, actually allowed to use mail order data for targeted advertising. In fact, the legal expert concluded that mail order data cannot be used for any purpose other than what they had originally been collected for, which was efficiently delivering mail items: “of course we get all these address data from business clients, because we [use that for delivery]. But actually, it was quite clear that [using them] was not possible: you cannot use these data for a different purpose” (Eva25, Legal). First, the legal expert reported to the strategy team that any data that can be used to directly or indirectly identify an individual should be categorized as “personally sensitive”. By law, personally sensitive data can only be used for the purposes for which they have originally been collected. Combined mail order data were personally sensitive, because they could indeed reveal information about individuals. The only way LogiCo might circumvent this issue was by abstracting the data to a level high enough such that individuals could not be identified. But the legal expert also advised that—even if abstracting and anonymizing the data were allowed—abstracted data might not lead to such useful insights anymore.

More importantly, the expert concluded that abstracting and anonymizing mail order data was not allowed, because the data were not owned by LogiCo in the first place. Rather, mail order data were owned by LogiCo’s business clients: the data had been uploaded by business clients under the premise that these data would be used for the purpose of efficiently delivering mail items. This ownership was specified in the metadata, that is, the contractual agreement between LogiCo and the business clients. If LogiCo wished to change or even (virtually) touch the data for purposes not directly linked to the delivery business, they would need to have explicit permission by the business clients.

25 All names are pseudonyms.
How Data Shaped the First Strategizing Move

<table>
<thead>
<tr>
<th>Mechanisms</th>
<th>Key Data Characteristics</th>
<th>Contextual conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Opening up choices and actions beyond the historical purpose of the data</strong></td>
<td><strong>Specificity of mail order descriptions</strong> Mail order data contain specific descriptions about the mail items that have been sent and the business clients that sent them. This allows analysts to gain insights into the type of mail that a household has received.</td>
<td><strong>External pressure</strong> The organization is pressured by a shrinking market and has been encouraged by external consultants to consider data-driven strategies based on mail order data.</td>
</tr>
<tr>
<td></td>
<td><strong>Time horizon of collected mail order data</strong> Mail order data have been collected and saved over a time span of several months. This allows analysts to gain insights into when a household has received certain mail items.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Relationality of mail orders from different clients</strong> Mail orders from different clients share an address field. Because of this relationality, mail orders from different clients can be readily combined. This allows analysts to gain insights into how many different mail items a household has received.</td>
<td></td>
</tr>
<tr>
<td><strong>Imposing historical decisions on current choices and actions</strong> Because of the nature of the mail order data, the organization is forced to only use them for the purposes for which they have originally been collected.</td>
<td><strong>Ownership of mail order data lies with clients</strong> Mail order data have been uploaded by business clients. Ownership of the mail order data therefore lies, by contract, with the business clients. Because of this, mail order data cannot be used for any purpose other than efficiently delivering mail items—at least, not without permission by the business clients.</td>
<td><strong>Strict regulation around sensitive data</strong> Regulation states that personally sensitive data may only be used for the purposes for which they have originally been collected.</td>
</tr>
<tr>
<td></td>
<td><strong>Sensitivity of combined mail order data</strong> Because analysts can deduce information about people living on an address, the data should, by law, be considered personally sensitive. Personally sensitive data may only be used for the purpose for which they have originally been collected.</td>
<td><strong>Strong organizational values</strong> The organization greatly values consumer privacy, as well as its relationships with business clients.</td>
</tr>
<tr>
<td><strong>Encouraging the inclusion of external actors in data-driven strategizing</strong> Because of the nature of the data, LogiCo would have to proactively include business clients in the strategizing process if they want to use mail order data for purposes other than what they have originally been collected for.</td>
<td><strong>Ownership of mail order data lies with clients</strong> Mail order data have been uploaded by business clients. Ownership of the mail order data therefore lies, by contract, with the business clients. Because of this, mail order data cannot be used for any purpose other than efficiently delivering mail items—at least, not without permission by the business clients.</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2 Mechanisms through which data shaped the first move.

In the first move, we see three different mechanisms through which the data influenced strategic choices and actions (see Table 3.2). First, we see that data initially opened up choices and actions beyond the historical purpose of the data. Because of the nature of mail order data, actors could think about offering products and services beyond their traditional delivery business, and beyond the purpose for which these data had initially been collected. Three characteristics of the data were especially important in this regard. First, the specificity of the descriptions: LogiCo had recorded quite specific information about each delivery order, including the delivery date, weight, target address, sender, and type of mail. Second, these data were collected over a time span of months, which allowed analysts to track when the different mail items had been delivered. Finally, the relationality of mail orders from different clients was important: data from different business clients could be combined because they all shared an address field. With access to detailed information about the types of mail
that addresses had received from various business clients over a period of time, analysts could gain insights into the personal preferences of consumers living on those addresses. This allowed LogiCo to explore strategic opportunities that reached beyond the traditional delivery business.

However, we also see that the data imposed historical decisions on current choices and actions. First, because the combined mail order data from different business clients could say something about the individual preferences of people living on an address, the data were also personally sensitive. By law, personally sensitive data could not be used for any purpose other than what they had originally been collected for. It was partly because of this personal sensitivity that LogiCo was forced to only use mail order data for their historical purpose. An even more prominent challenge, however, was the fact LogiCo did not have ownership over the data. Ownership of the data was specified in the contracts between LogiCo and the business clients. Because of this, the data could not be used for any purpose other than what they had originally been collected for—at least, not without specific permission by all stakeholders.

Because ownership of the data was in the hands of business clients, LogiCo was also encouraged to include external actors in the process of data-driven strategizing. They were encouraged to talk to business clients and see if there was any way in which data ownership could be negotiated, or perhaps to find ways in which both LogiCo and the business clients could benefit from shared ownership. Near the end of 2012, however, a new opportunity presented itself that marked the beginning of a second strategizing move.

3.4.2 Second Strategizing Move: Selling External Data and Providing Analytical Insights

Near the end of 2012, LogiCo was approached by DataCo, a small data-driven startup that had been obtaining data from public sources (e.g., the Agency for Statistics); had been scraping data from several websites (e.g., auction sites and sites offering real estate information), and had bought data from parties such as the Public Land registry. DataCo’s owner approached LogiCo after hearing about their data-driven initiatives on the radio and thinking that the data collected by DataCo might be useful to LogiCo. After evaluations and negotiations, LogiCo officially acquired DataCo in 2013.
Table 3.3 Summary of the second move.

DataCo collected data about geo-demographics and real estates, often recorded at the level of addresses, streets, or postal codes. LogiCo’s strategy team believed that these data would fit well with their postal service business and their intentions to offer targeted advertising: “We didn’t know anything about addresses. And [DataCo] had built a whole database with all kinds of information about an address” (Rick, Strategy team member). The update frequency of these data varied between daily, weekly, monthly, or yearly—this also depended on the original source from which the data were being collected. For example, data from the Agency of Statistics were generally updated once a year. After the acquisition of DataCo, LogiCo’s representatives emphasized towards their clients that mail order data and the data from DataCo were physically separated, and that they greatly cared about consumers’ privacy. The data by DataCo had been legally retrieved from other sources; DataCo had been collecting these data for many years and the owner had ensured that this happened in a legally responsible way. If needed, analysts aggregated data to a higher level such that individuals could not be traced.

Now that they had access to the data collected by DataCo, LogiCo’s development teams could begin developing data-driven products around targeted advertising. For one of these products, sales team members would bring lists of data attributes to meetings with their business clients. Together, the sales team member and the business client would select combinations of data attributes that were representative for the business client’s target groups. For example, a business client might be looking for addresses of young people with high incomes. Based on such selection criteria, analysts would filter out, from
their databases, those addresses of young people with a high income. The resulting list of addresses could be sold at a price per address, or at a combined price for both the addresses and the delivery of mail items to those addresses. If business clients were not exactly sure what type of customers they were currently targeting, they could also submit a list of addresses of their existing customers. Based on the data in LogiCo’s database, analysts could then see how these customers scored on several attributes. For example, the analyst might find that a client’s existing customer base consisted mostly of elderly people with gardens. LogiCo would sell a report with these kinds of insights on request.

In practice, sales team members found it difficult to convince business clients of the value of the data and data-driven products. Traditionally, sales teams had been selling large volumes of mail at a price lower than the price offered by their competitors. Now, they had to change their mindsets and engage in a type of consultative selling, that is, sit together with business clients and help them think about how they can reach their prospective customers. This proved to be quite difficult for them. Part of the issue was that it took effort for business clients and sales team members to select and combine data attributes in such a way that they would become representative for the business client’s target group. For example, they would have to decide on such things as how old these people are, whether they own children, whether they own a garden, and what more. Deciding on all the different attributes was perceived as challenging: “If you talk about all those attributes people quickly get tired. It is difficult for our sales people, and often also for an average marketing manager [..], it’s much nicer to talk about profiles than when they have to talk about those attributes” (Jake, Business Intelligence). Consequently, data-driven products were hardly being sold and sales teams tended to direct their attention at selling mail items.
Table 3.4 Mechanisms through which data shaped the second move.

We can see two mechanisms through which data shaped the second strategizing move. First, DataCo’s data opened up choices and actions beyond the historical purpose of the data for LogiCo. A number of characteristics were important in this regard. For one, data were accessible both from a technical and contextual perspective: analysts could buy, obtain, and scrape data from several external sources. These data could be used for purposes other than what data providers had initially collected the data for (e.g., for real estate), because ownership was absent or could be negotiated. Additionally, the data from different sources all contained (parts of) an address field. Because of this relationality, analysts could quite easily combine data from several sources.

As sales team members began offering data-driven products, however, the data also hindered business clients to become engaged in data-driven strategizing. Although sales team members and business clients did not work directly with the digital data records, they used a checklist that represented the different data attributes. These attributes were rather granular and could be
combined in many different ways. As described in the narrative, business clients and sales team members found it difficult to combine these granular data attributes in such a way that the combination of attributes would represent the business client’s target groups. To overcome this challenge, analysts thought of new ways to engage business clients and more actively involve them in value creation activities, thereby initiating a third strategizing move.

3.4.3 Third Strategizing Move: Offering Profiles and a Tool to Engage Business Clients

In the third strategizing move, development team members actively tried to find ways in which they could engage business clients in value creation activities, and allow them to make sense of the data and data-driven products.

<table>
<thead>
<tr>
<th>Move 3: Offering Profiles and a Tool to Engage Business Clients</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong> To try to engage business clients and allow them to make sense of the data, an analyst develops a limited number of profiles or “personas” based on combinations of different data attributes. Additionally, LogiCo outsources the development of a tool, which allows business clients to interact with the data more directly.</td>
</tr>
<tr>
<td><strong>Data-Related Practices</strong></td>
</tr>
<tr>
<td>- Analysts collect data from external sources by scraping, buying, or otherwise obtaining them.</td>
</tr>
<tr>
<td>- Analysts combine data from different sources based on the address field.</td>
</tr>
<tr>
<td>- Analysts aggregate data that may be sensitive to a higher level of abstraction.</td>
</tr>
<tr>
<td>- Analysts translate higher-level data (e.g., street or postal code level) to lower-levels (e.g., address level).</td>
</tr>
<tr>
<td>- Analysts categorize data attributes to construct profiles representing household types.</td>
</tr>
<tr>
<td>- Business clients and sales team members select relevant profiles and data attributes through a tool, after which addresses are automatically filtered.</td>
</tr>
<tr>
<td>- Analysts compare business clients’ address lists against their own database to validate the address list.</td>
</tr>
<tr>
<td>- Analysts compare business clients’ address lists against their own database to gain insights into how business clients’ existing customers score on different attributes.</td>
</tr>
<tr>
<td><strong>Data:</strong> Profiles based on Data from DataCo</td>
</tr>
<tr>
<td>Data are now offered to business clients in the form of profiles. Thus, the data that business clients and sales team members interact with (through an interface) are more abstract than the original data collected by DataCo.</td>
</tr>
<tr>
<td><strong>Challenges:</strong> The perceived value of data and data-driven products is limited.</td>
</tr>
</tbody>
</table>

Table 3.5 Summary of the third move.

To help sales team members in their conversations with business clients and convince clients of the value of data-driven products, an analyst categorized data attributes into 14 higher-level profiles. Examples of profiles are “young urban professionals”, “traditional starter”, and “young ambitious family”. To construct these profiles, the analyst asked an expert in household profiling for advice and gained inspiration from profiles that competitors had been offering. Next to developing the profiles, development team members outsourced the development of a tool through which business clients themselves could select attributes to arrive at a useful list of addresses. The tool, that went live in the Fall of 2015, was designed by the development team, but built by an external development company. It was built to make life easier for LogiCo’s analysts, sales team members, and business clients: analysts would no longer have to do selections and analyses manually—which up until then had been tedious and time-consuming tasks for them—and sales team members could use the
tool in their conversations with business clients as a means to explain data-driven products. Clients could use the tool to personally select data profiles and attributes and buy a list of addresses, without relying on sales team members.

Initially, it seemed that having the profiles at least made it easier for business clients to connect with the data and understand how they may be relevant to their context: “if you say, I’m looking for an ‘active senior citizen’, yes that is exactly my target group, that was easier” (Quinten, Sales & Marketing). Unfortunately, actors ran into additional problems as they were working with the data in practice. First, the profiles were relatively broad (there were only 14 of them), and business clients wished to sometimes make more specific selections within a profile (e.g., elderly people, but only those of ages between 60 and 65). Second, business clients often wished to select attributes that were not at all included in the profiles or in the data that LogiCo had collected. Business clients would, for example, request the names of the individuals living on addresses, or desired attributes that were specific to their business (e.g., whether households own a pet). LogiCo then needed to contact other data providers to see if they could gain access to such attributes. Third, LogiCo found it difficult to prove to business clients that targeted advertising was actually more effective than undirected advertising, and would lead to increased conversion. Not only was it difficult to measure the effect of physical advertisements, but stakeholders also experienced issues with the data. In some cases, it appeared that the data had been outdated (e.g., when mail items were returned because residents had moved). In other cases, actors felt that the level of observed granularity may not have been granular enough for effectively targeting potential customers: “[..] in my case, I live in a street with all semi-detached houses. We all have a garden, we all have a roof. So actually, it’s still a little uniformly. But the data about people living in that building, ehm yeah those came from public sources. Those said something about a street or region” (Robert, Sales). As they later reflected on the issue, some interviewees generally questioned whether address-level data—as opposed to data about individuals—had been detailed enough to target individuals living on an address.

Gradually, strategy team members, sales team members, and business clients began to perceive the value of data and data-driven products as limited. Addresses were hardly being sold, data-driven products did not result in as much revenue as initially expected, and resources to improve and invest in the data had become limited. The CIO had become seriously ill and the CCO was replaced—these had been two of the main advocates of data-driven initiatives. The remaining strategy team members decided to allocate most of the resources to initiatives that directly focused on gaining more market share in the traditional postal market. By the fall of 2015, development team members were convinced that any data-driven initiative should always support the existing...
business, and that the focus should be on selling more mail volumes. Near the end of 2015, a new CEO and CCO were appointed. The new CEO announced a strategic reorganization that marked the ending of the third strategizing move.

**How Data Shaped the Third Strategizing Move**

<table>
<thead>
<tr>
<th>Mechanisms</th>
<th>Key Data Characteristics</th>
<th>Contextual conditions</th>
</tr>
</thead>
</table>
| Hindering external actors to engage in data-driven strategizing | Abstract nature of data profiles  
Data are presented in the form of profiles, but business clients wish to make more specific selections. The profiles may be too abstract for business clients to become engaged. | Limited user skills  
Sales team members and business clients have limited skills to understand and work with data. LogiCo does not have resources to invest in developing skills or capabilities. |
| Imposing historical decisions on current choices and actions | Misalignment between timeliness and current strategy  
The timeliness of the data from external sources is in line with the original purpose for which they have been collected. As LogiCo uses them for a different purpose, some data elements appear to be out-of-date, e.g., when mail items are returned because someone has moved houses.  
Misalignment between observed level of granularity and current strategy  
Data from external sources have been collected at the level of addresses, streets, or postal codes. As LogiCo uses them to target households, some of these data may not be granular enough.  
Address-level data as a result of relationality  
Data from different sources all share (parts of) an address field. Because of this relationality, data from different external sources can be readily combined. But because of this relationality, actors may be less inclined to find data beyond the level of addresses. | Limited resources to invest in data  
LogiCo has limited resources to invest in improving the existing data and acquiring new data.  
Strong business logics  
LogiCo had always delivered mail items to addresses on behalf of business clients. Actors have strong mindsets towards this traditional business. |
| Encouraging the inclusion of external actors in data-driven strategizing | Diversity as a key success characteristic  
The more data attributes, the better the data lend themselves for targeted advertising. But to gain access to many different attributes requires collaboration and networking with several data providers. |  |

Table 3.6 Mechanisms through which data shaped the third move.

In the third move, we see that the data still hindered external actors to become engaged in data-driven strategizing. While in the second move the data hindered external actors because the attributes were too granular for them to make sense of the data, data presented in the form of profiles were now considered too abstract for business clients to see how these data may be relevant to their specific context and engage in value creation processes. While the profiles made more sense to the business clients, business clients now sometimes wished to make more specific selections which they could not make based on the profiles and the attributes offered through the tool.

We could also see several ways in which the data imposed historical decisions on current choices and actions. First, the timeliness of the data proved to be a serious issue for LogiCo; some of the data that LogiCo relied on appeared to be out-of-date as mail items were being returned. Yet, this could
not always simply be solved by more frequently collecting or scraping the data. Rather, the update frequency of the data was determined by those who had collected the data in the first place, and was in line with the original purpose for which the data had been collected. Here, one interviewee provided the example of data collected from a real estate website: those data are only updated when a real estate is put up for sale. Historically, those data were not meant to be continuously updated, and the data provider may have no logical reason for doing so. As actors at LogiCo worked with data, they were confronted with a misalignment between the historical timeliness that was inscribed in the data, and the purpose for which LogiCo aimed to use them.

Similarly, the level of granularity at which the data had traditionally been observed and recorded by data providers became an issue for LogiCo. As described, some of the data collected from external sources had been recorded at the level of streets or even postal codes, rather than address level. Indeed, we may question to what extent the average age in a street or postal code area is representative for the (average) age of people living on one address. Yet, it had made sense for those who originally collected these data, such as the National Bureau of Statistics, to collect data at this level of granularity. As LogiCo’s people worked with the data from external sources in practice, they were confronted with a misalignment between the level of analysis inscribed in the data and the purpose for which LogiCo aimed to use them.

We identified another, slightly less obvious way in which data imposed historical decisions on current choices and actions. As mentioned in the narrative, individuals at some point began to question whether address-level data in general was detailed enough for the purpose of targeted advertising, and why they had not collected, for example, data about individuals. Yet, it had made perfect sense for actors inside LogiCo to focus on collecting address-level data: LogiCo had always delivered mail items for business clients to addresses, and the data that actors had worked with over the past years were structured around addresses. Strategy-team members immediately saw how data on address level could be linked to their traditional data, which they relied on to execute their traditional strategy. As LogiCo collected more and more data, over time, the address-field consistently became the most prominent denominator based on which data from different sources could be combined. This illustrates that Logi-Co’s traditional address-level focus had become inscribed in the data that they consistently used as the basis for combining other data. Thus, the relationality of address-level data with other address-level data allowed LogiCo to combine data from different data sources, but also diverted them from looking at other, perhaps more granular data.

Finally, we see that, because the diversity of the data had become a key characteristic, LogiCo was encouraged to include external actors in the
The process of data-driven strategizing. The diversity of data was a very important characteristic enabling the use of data for targeted advertising: the more data attributes, the better the data could be used to describe households, and the more business clients might be served. Business clients also began requesting additional attributes (e.g., whether households own pets). However, not all of these data elements could be collected, maintained, and updated in-house: this would require large investments in technologies and expert capabilities, which LogiCo did not have resources for. Nor was there an economic justification for collecting and storing data each time a business client requested new data attributes. Thus, because diversity had become a key characteristic, collaborating and partnering with external data brokers became a logical next step for LogiCo, which is also something that they relied heavily on in the fourth move.

3.4.4 Fourth Strategizing Move: Connecting Business Clients and a Data Broker for Targeted Advertising

When the CIO and many of those involved in developing data-driven products officially left the organization between mid-2015 and the spring of 2016, data-driven products were placed within the marketing department, under the care of a product marketer. Given that several stakeholders experienced issues with the data, the product marketer who had become in charge of the data-driven products felt that they would have to consider alternatives and find new data sources if they were going to continue offering the data-driven products.

Mid-2016, a few business clients indicated that they were still interested in the idea of more targeted advertising. One of the sales team members, who had been discouraged to look at LogiCo’s own data because of the limited perceived value, then decided to look for data elsewhere. In the search for new data sources, the sales team member came in contact with a data broker that had access to data from many different sources, especially about addresses. The sales team member brought a business client in contact with this data broker and, together, the sales team member, the data broker, and the business client selected relevant addresses for the business client’s mailing. Because this case was a success, the sales team member presented the results to LogiCo’s strategy team. It was then decided that LogiCo would act as an intermediary between data providers and LogiCo’s business clients.
The tool that LogiCo had developed earlier was taken offline in January 2017. Since then, LogiCo relied on the data broker on a pay-per-use basis: “So we don’t have to maintain those data, we don’t have to buy them. Actually, we rent them, only those data for that single customer, for that single moment. It makes it much more controllable, much leaner. And it leaves us free to keep looking for other data providers” (Quinten, Sales & Marketing). The data broker created a standardized checklist with data attributes, which LogiCo’s sales team members could bring to their conversations with business clients. Sales team members offered business clients a combined price for both data and delivery. If the business client wanted to select characteristics other than those offered on the checklist, the sales team member would contact the data broker to see what they could do: “we know the way in ‘Data-Land’. You see it happening at [LogiCo] […] The most exotic selections are proposed: ‘I need everyone with this income, driving these types of cars’, […] everything is being asked for [by business clients]” (Tim, Data Broker). Business clients could request any characteristic as long as the data broker knew where to find data. The data broker could also give business clients insights into what types of households they were targeting at the time, or help to validate a business client’s mailing list. In some cases, representatives of the broker would work more closely with business clients and sales team members to help them arrive at a relevant set of addresses, for example, by joining them in their conversations or giving them advice. In such cases, they acted as a data partner rather than only a supplier.

3.5 Discussion

By bringing data to the forefront, we were able to gain insights into how data can shape the process of data-driven strategizing. Our findings illustrate that different characteristics of data may not only enable and constrain certain actions, but may also encourage actors and organizations to move into certain strategic
directions. Specifically, we identified four different mechanisms through which data can shape strategic processes and outcomes, that is, by: 1) opening up choices and actions beyond the historical purpose of the data, 2) imposing historical decisions on current choices and actions, 3) encouraging the inclusion of external actors in data-driven strategizing, and 4) hindering external actors to engage in data-driven strategizing. Considering the four mechanisms and the ways in which they have been enacted in our case, we argue that they essentially reflect two tensions. Below, we will first explicate these tensions, before discussing the theoretical and practical implications of the findings.

3.5.1 Enabling Exploration vs. Imposing Historical Decisions

On the one hand, we empirically affirmed that data may open up strategic choices and actions beyond the historical purpose of the data, which seems to be in line with much of the literature that argues that data can be useful beyond the purpose for which they had initially been collected, thereby allowing for strategic exploration (e.g., Aaltonen & Tempini, 2014; Woerner & Wixom, 2015; Constantiou & Kallinikos, 2015; Zeng & Glaister, 2017). We saw in our case that characteristics of the data such as specificity, relationality, diversity, and the time horizon can significantly contribute to opening up a space of opportunities for organizations to create new products and services. Especially the relationality of data consistently played an important role in the different moves, because it made it possible for different data elements to be combined and synthesized. The more data from different sources could be combined on address-level, the better the data could be used to describe households. This “interconnectivity” of data—defined as the “possibility to synthesize data” (Günther et al., 2017, p. 201)—has previously been identified as one of the key characteristics of “big data” that allows for the exploration of new strategic opportunities (e.g., Malgonde & Bhattacherjee, 2014; Zeng & Glaister, 2017).

However, the case also illustrates that data may impose historical decisions on actors’ current choices and actions. Specifically, we saw that past choices and actions can become inscribed in the characteristics of the data: as actors work with data over time, the characteristics of these data come to reflect the different choices that have been made by these actors. For example, the level of granularity or the timeliness of the data may align with the historical purpose for which the data have originally been collected. When organizations then use such data in the process of strategizing, they may be confronted with misalignments between the historical choices and actions that are reflected in the characteristics, and the current purpose for which actors aim to use the data. In the case, for example, we saw that actors were confronted with a misalignment between the level of granularity and the timeliness of the data required for the original purpose, and the level of granularity and timeliness
required for the purpose of targeting prospective customers.

Interestingly, we also found that the same characteristics that contribute to opening up strategic opportunities for organizations, may also cause organizations to stay close to the original purpose of the data. In the case, for example, we saw that while the specificity, relationality, and time horizon of mail order data allowed analysts to gain insights that might be useful for offering targeted advertising, it was also because of these characteristics that the mail order data should be considered personally sensitive. And data that are personally sensitive, as LogiCo learned, may only be used for the original purpose for which they have been collected. This suggests that in contexts where there are strict rules and regulation around personally sensitive data, and where organizations do care about the social implications of data analytics, characteristics like specificity, relationality and timeliness may actually form a double-edged sword. The more detailed the information that is inscribed in the data, and the easier the data can be combined and collected over periods of time, the higher their technological “quality” might be; but the more organizations in such contexts may be forced to stay close to the original purpose for which the data have been collected.

We also found that the same relationality that enables actors to combine data and gain insights from their synthesis, may also prevent actors from looking beyond traditional choices and actions. In LogiCo’s case, the relationality of data consistently allowed for the combination of data from different sources on address level, and contributed to the fact that insights about addresses could be gained from the synthesis of the data. Paradoxically, however—as explained in our reflection on the role of data in the third move—this same relationality may also have prevented LogiCo’s actors from considering data beyond the level of addresses. LogiCo’s stakeholders became primarily concerned with combining data on address level, because they had always been delivering mail to addresses, and because such new data could so easily be combined with the existing address-level data. This suggests that when actors have relatively strong mindsets towards the traditional business logics, and when the data also reflect choices that are in line with the traditional business strategy, then the relationality of data may only enforce such traditional ways of thinking.

3.5.2 Encouraging Collaboration vs. Hindering Stakeholder Engagement

By unpacking the characteristics of data in our case, we were also able to see how, because of the nature of the data, organizations may be encouraged to collaborate with different external stakeholders. For instance, we found that when organizations do not have ownership over the data, managers will have to negotiate ownership and may be forced to include the data owners in the
process of data-driven strategizing. In the case of LogiCo, ownership of mail order data was in the hands of business clients, and LogiCo may have collaborated with them to see if there are ways in which they can use these data to the benefit of multiple stakeholders. We also found that when the diversity of the data is a key success characteristic, organizations may have to reach out to external data providers. In LogiCo’s case, business clients increasingly began requesting many different attributes that were specific to their business and their specific context. Yet, there was no economic justification for collecting data every time a business client made such a request; therefore, LogiCo relied on partners who could provide such data. In similar situations where the diversity of data is especially important, organizations may have to rely on connections with external stakeholders to gain access to many different data attributes (Schuritz et al., 2017).

Our findings also indicate that data may hinder external stakeholders to become engaged in data-driven strategizing. We saw that LogiCo especially struggled to find a balance between how granular or abstract the data should be, to allow business clients to be involved in value creation processes. While business clients found it difficult to combine raw attributes (such as age, income, and education) in meaningful ways, they also considered higher-level profiles to be too abstract and wished to make more specific selections. Business clients were essentially trying to see how the data could become relevant to their organization and their specific context. Offering a tool through which business clients could access and interact with the data did not solve this issue—business clients would often still call LogiCo’s sales team members for help. Only in the last move were different stakeholders more effectively able to make sense of the data, when business clients and sales team members would work together with a data broker to see what data they needed and how these data could be used to effectively target prospective customers. This affirms that in order for organizations to create value with data, data need to be “contextualized” (Zeng & Glaister, 2017). An important element of contextualizing data then is to make sure that different stakeholders can “assign meaning” to the data and that there is a shared understand of what the data represent (Zeng & Glaister, 2017, p. 20; Jones, 2018). Additionally, different stakeholders may need to work together to see how data are relevant to a specific context, or might be changed in such a way that they can become relevant.

3.5.3 Implications for Theory

Our findings have a number of implications for theory. First, we contribute to the literature that aims to understand how organizations may successfully leverage data as strategic resources. As previously argued, this emerging literature stream has so far been pushing data to the background. By conducting a case
study, unpacking different characteristics of data inside the context in which they were used, and examining how these data influenced strategic choices and actions over time, we were able to identify nuanced mechanisms through which data can shape strategic processes and outcomes. Our findings highlight the need for future studies to treat data as “actors” that have distinct characteristics, can play different roles in different contexts, and of which the role may constantly change over time. By doing so, scholars may gain complementary insights that help us understand why organizations may still struggle to realize the expected benefits from adopting data-related technologies (Gartner, 2018-2; Ransbotham et al., 2016).

Our findings also have implications for a stream of literature that promotes the various opportunities associated with “big data”. Specifically, our findings cast criticism on one of the major assumptions associated with big data, which is that data are “unbounded” and infinitely open to be explored with many emergent value propositions (e.g., Aaltonen & Tempini, 2014; Constantinou & Kallinikos, 2015). Our findings empirically illustrate that historical decisions may be inscribed in the data, and that these decisions may be imposed on current and future choices and actions. Even characteristics such as the level of detail and relationality, which have almost consistently been considered as characteristics that open up numerous opportunities for organizations (e.g., Aaltonen & Tempini, 2014; Lycett, 2013; Yoo, 2015; Zeng & Glaister, 2017), may also lead organizations to stay close to the traditional purpose of the data. Future research may look at how actors can potentially “break free” from the historical choices and actions that are inscribed in the data in their pursuit of innovation.

Our findings question another assumption that underlies much of the literature on big data (Günther et al., 2017, p. 201), which is that data are inherently portable and can be readily transferred to and used in many different contexts (e.g., Lycett, 2013; Tallon et al., 2013-14). We found two reasons for why data may not be as portable as is often suggested. First, we found that even when data are accessible from a technological perspective, social characteristics such as the ownership of the data may pose serious challenges. We invite scholars to also consider these non-technical characteristics of data and study their influence on organizing processes. Additionally, our findings illustrate that stakeholders may need help to be able to assign meaning to the data and see how the data are relevant to their specific contexts. Rather than assuming that data can always readily be transferred and used across different contexts, we need to understand how data can be effectively “contextualized”. Future research is needed to understand how organizations can ensure that different stakeholders are able to make sense of the data and understand how the data can be(come) relevant to their specific contexts (Zeng & Glaister, 2017). Generally, future research may look at how the contextualization of data across
different contexts can be supported.

Finally, we would like to stress the value of taking a process view when studying the role of data in organizing processes. We argued beforehand that data are extremely malleable resources that may continuously change as actors work with them in practice (Ekbia, 2009; Kallinikos et al., 2013), and that new opportunities and challenges may emerge over time as actors interact with the data (Constantiou & Kallinikos, 2015; Aaltonen & Tempini, 2014). Our findings illustrate that in addition, past choices and actions may be inscribed in the data (or metadata), that subsequently impact current and future choices and actions. Scholars can capture this historical situatedness and its influence on organizing processes by adopting a process lens and carefully assessing: where did the data come from; how did they change over time; what was the historical purposes for which the data had been collected and changed, and how is this reflected in the data?

3.5.4 Implications for Practice

Our findings also have several implications for practice. First, practitioners need to be aware that not all data are the same, and that data with different characteristics may differently influence strategic processes and outcomes. We explicated a number of characteristics—not just “data quality” characteristics—that influenced strategic choices and actions in the case of LogiCo (e.g., granularity, relationality, ownership). Managers may see if and how similar characteristics play a role in their specific context. Still, rather than assuming a fixed list of characteristics, managers should be open to see how any characteristics of data may be relevant in their specific context. For example, in their specific case, the volume, compactness, or reproducibility may be relevant[26]. Importantly, in order to be able to grasp the different characteristics of the data and understand how these data may influence strategic choices and actions, the different actors involved in data-driven strategizing, including top managers, will need to have some affinity with the data (Lee et al., 2014).

Second, practitioners can learn from our study that it is not only important to know what data you have, and how this may be relevant to the envisioned strategy, but also how this data has developed over time and what past choices may be inscribed in the data. In other words, managers should pay close attention to where the data came from. When the data come from outside of the organization, practitioners need to understand what these data had originally been collected for and how this purpose is reflected in the different characteristics of the data. But even when the data come from within the boundaries of the organization, managers need to be aware of how traditional ways of thinking may be inscribed in the data.

[26] See Wang and Strong (1996) for a list of 179 potentially relevant characteristics.
Third, managers can learn from our study that because of the nature of the data, they may be encouraged to closely collaborate with external stakeholders. For instance, organizations may have to negotiate ownership over the data, and may be encouraged to partner with other organizations to gain access to (complementary) data sources. In doing so, managers need to carefully assess what their contribution will be. This is especially important when relying on external stakeholders to gain access to the data: managers will need to evaluate who else has access to these data—this is also part of knowing where the data came from—and what the organization can offer as a unique and durable contribution (Kiron et al., 2014). For example, LogiCo in the end came to fully rely on the data and data-related capabilities offered by a data broker, such that they themselves could effectively focus on the actual delivery of mail items. This was an interesting move, given that they initially embarked on their data-driven journey because the postal market was shrinking.

Finally, managers can learn from our case that data-driven value realization is often not a straightforward process, rather, it is a journey: “finding value from data is much more a process of cultivation than it is one of extraction or refinement” (Thorp, 2012). Similar to LogiCo, organizations may find themselves continuously revisiting their data or their data-driven strategies. Even with the best data scientists, the most sophisticated systems, and the most experienced people, you may be taken by surprise. Perhaps the key takeaway from our paper that practitioners should take with them on their data-driven journey is: mind your data.

### 3.5.5 Boundary Conditions and Avenues for Future Research

The findings presented in this study are based on a single case study. As argued in the theoretical background, however, data may play different roles across different contexts. Thus, future research may effectively examine the role of data in non-similar cases and see if they can find conflicting or complementary mechanisms to the ones that we found. Our case, for example, concerned an organization that was active in a rather traditional market. Future research may focus on examining the role of data in the context of organizations that are inherently more “digital” (Bharadwaj et al., 2013), such as online marketing agencies. Similarly, while the postal service market in Europe is quite heavily regulated, future studies may examine how data shape the process of data-driven strategizing in markets that are more open, to see to what extent characteristics such as “ownership” and “personal sensitivity” influence strategic choices and actions. Studying the role of data in contexts in which organizations more intensively collaborate with other stakeholders would also be especially interesting, and may particularly lead to insights regarding how organizations can effectively “contextualize” data across different contexts. Finally,
future research may examine the role of data in contexts in which organizations more heavily rely on sophisticated algorithms and “Artificial Intelligence”. This is especially relevant in light of recent debates that argue that big data “actively shapes the world. And it does it through algorithms” (Yoo, 2015, p. 64).

ACKNOWLEDGEMENTS

We would like to thank Sue Newell for her valuable suggestions and engagement in compelling discussions. We also gratefully acknowledge the European postal service organization for granting us access to their organization and providing the financial support for this research. Finally, we would like to acknowledge all members of the KIN Research group and the reviewers and participants attending our presentations at several conferences (ECIS DC 2017, ICIS 2017, PROS 2018, AOM specialized conference 2018, and AOM 2018).