6-18-2022

Inside a Data Science Team: Data Crafting in Generating Strategic Value from Analytics

Marta Stelmaszak
Portland State University, stmarta@pdx.edu

INSIDE A DATA SCIENCE TEAM: DATA CRAFTING IN GENERATING STRATEGIC VALUE FROM ANALYTICS

Research Paper

Marta Stelmaszak, Portland State University, Portland, USA, stmarta@pdx.edu

Abstract

Current research agrees that the value of data lies in analytics that generate valuable insights for strategic purposes. However, little is known about how these insights are derived by data scientists. This research reports on the work of an embedded data science team at an organization striving to use people analytics to improve its strategic human resource management. We find that to create strategically valuable analytics, data scientists engage in data crafting, an approach to data science work that relies on broadcasting the potential value of data science towards the organization, cultivating a shared vision of value within the team, and creating value-adding data products with organizational customers. To do so, the team requires appropriate positioning and autonomy within the organization. Our findings have implications on understanding the role of data science teams and organizational data with respect to strategy, and practical insights for realizing strategic value from analytics.

Keywords: Data science, data work, strategic value, analytics, craft.

1 Introduction

As organizations invest resources into establishing and developing data science teams charged with conducting analytics (Davenport, Harris and Morrison, 2010; Davenport, 2018), both practitioners and researchers alike investigate how to ensure that the outputs that these teams produce generate strategic value. Such value may come in the form of improved performance, better competitiveness, faster decisions, or enhanced business processes (Wamba *et al.*, 2017; Ghasemaghaei, Ebrahimi and Hassanein, 2018; Grover *et al.*, 2018), but the pathway from data to strategic value through analytics is far from being clear. In fact, there have been cases where data were found to lead to strategic failure instead (Aversa, Cabantous and Haefliger, 2018; Joshi *et al.*, 2021). Against the still predominant view that assumes data science to be a technical, nearly mechanistic endeavor whereby data scientists play only an instrumental role (Davenport and Harris, 2017), information systems (IS) researchers begin to study data work in more detail. For example, Joshi (2020) investigates how data scientists rely on both subjectivity and objectivity in the production of information from data, van den Broek and colleagues (2021) show that their work relies on a hybrid practice combining machine learning and domain expertise, and Parmiggiani and colleagues (2022) elucidate the lack of stability and pervasive bidirectionality of data science work.

Research in this area so far focuses mostly on what data scientists do as they produce analytics outputs, but we do not know what mechanisms ensure that their work has the potential to contribute to strategic value. Data science work relies on deriving new information, uncovering patterns and anomalies, obtaining hidden knowledge, or deriving actionable insights (Zhang, Wang and Pauleen, 2017; Chiang *et al.*, 2018; Castillo *et al.*, 2021)–in short, on generating new organizational knowledge–out of organizational data. The nature of this work, especially its highly iterative, exploratory, and non-linear character (Patel *et al.*, 2008; Kery and Myers, 2017; Rule, Tabard and Hollan, 2018), and the sense of professional identity (Vaast and Pinsonneault, 2021) and occupational community (Avnoon, 2021) shared by data scientists suggest that craft, "a humanist approach to work that prioritizes human engagement over machine control" (Kroezen *et al.*, 2021, p. 502), may be a plausible lens to study *how data scientists create analytics outputs that contribute to strategic value*.

In response to this question, we conducted a single-case study (Yin, 2018) of a data science team embedded in a people analytics department at the human resources function in a large, multinational technology company. Drawing on interviews as a data collection method, and underpinned by grounded theory approaches to data analysis, we observed the ways in which the team ensured the strategic value of its work at a time of structural and managerial change. Because of its embedded nature in a cost-generating support function, the team was aware of increased pressures on ensuring the value of its work, and therefore keenly engaged in practices leading to generating strategic value of analytics. We found three mechanisms of ensuring that analytics outputs contributed to strategic value: *broadcasting the potential value of data science* towards the organization by establishing visibility and finding customers, *cultivating a shared vision of value* within the team by building communality, aligning with strategy, and engaging in continuing development, and *creating value-adding data products* by identifying the scope of data, productizing data, and co-designing products with customers. These mechanisms are underpinned by craft skills, attitudes, and conditions, and together they are constitutive of what we term *data crafting*, that is an approach to creating strategically valuable analytics outputs underpinned by the positioning and autonomy of the data science team.

We contribute to extant research by proposing a model of how data science teams can create strategically valuable outcomes from analytics which requires data science to be positioned within strategy-making. We show that the new forms of data that permeate organizations and are seen as potentially valuable rely on more human involvement from even more skilled and specialized professionals than previous, structured data, and that they need to be skillfully crafted into organizationally relevant data products. Finally, we identify positioning and autonomy as two craft conditions which may be used by organizations in practice to foster data crafting.

2 Background literature

2.1 Data science and strategic value

Various new forms of data that organizations have turned their interest to, such as social media data (Alaimo and Kallinikos, 2017; Castillo *et al.*, 2021), transaction data (Lehrer *et al.*, 2018; Leidner and Tona, 2021), trace data (Berente, Seidel and Safadi, 2019) or sensor data (Monteiro and Parmiggiani, 2019; Chen, Preston and Swink, 2021), are seen as potential sources of value. It has been argued that it is not only, and perhaps not mainly, the volume, velocity or variety of these data (Constantiou and Kallinikos, 2015; Yoo, 2015; Abbasi, Sarker and Chiang, 2016; Baesens *et al.*, 2016) that hold the biggest promise, but also other qualities that make them stand out against the previous forms of structured, purposefully collected forms of organizational data. Notably, data that are unstructured, unpurposeful, heterogeneous, and performative (Kallinikos, Aaltonen and Marton, 2013; Constantiou and Kallinikos, 2015; Yoo, 2015; Grover *et al.*, 2020) are often seen as enabling organizations to tap into new strategic value. Because of the different nature of these data, they are less accessible to organizational members, who are not equipped with the right tools and skills to work with data that do not come neatly structured and organized in spreadsheets. Indeed, this new type of data often challenges existing data management practices (Chen, Chiang and Storey, 2012; Bhimani, 2015; Woerner and Wixom, 2015; Chen, Preston and Swink, 2021) and ways of working (Constantiou and Kallinikos, 2015; Abbasi, Sarker and Chiang, 2016; Chiang *et al.*, 2018). For this reason, it has been acknowledged that the key to deriving value from data lies in their suitable preparation, processing, analysis, and visualization, often referred to as analytics, or with a more encompassing term data science. Data science in organizations focuses on extracting informative patterns from data (Saar-Tsechansky, 2015) which allow the understanding of real-world phenomena through the analysis of data with the goal of deriving business outcomes (Ghasemaghaei, Ebrahimi and Hassanein, 2018; Müller, Fay and vom Brocke, 2018; Cybulski and Scheepers, 2021). This work is done by data scientists who possess the right skills and abilities to work with the data (Avnoon, 2021; Vaast and Pinsonneault, 2021).

Data science work then focuses on generating value out of data through deriving new information (Teo, Nishant and Koh, 2016; Someh, Shanks and Davern, 2019; Castillo *et al.*, 2021; Pan *et al.*, 2021), uncovering patterns and anomalies (Chen, Preston and Swink, 2015; Pan *et al.*, 2021), obtaining hidden knowledge (Chiang *et al.*, 2018), deriving actionable insights (Zhang, Wang and Pauleen, 2017; Chen, Preston and Swink, 2021; Rana *et al.*, 2021), or improving the understanding of markets (Chen, Chiang and Storey, 2012; Sharma, Mithas and Kankanhalli, 2014). Just "generating and collecting social data might not be enough to use information efficiently" but monitoring, analyzing and identifying patterns in these data is critical for generating strategic value (Castillo *et al.*, 2021, p. 1473), and data scientists are crucial in the process of transforming data into valuable insights (Ransbotham *et al.*, 2016; Mikalef and Krogstie, 2020). Data science outputs can contribute to strategic value through enhancing business operations (Davenport, Barth and Bean, 2012), improving performance (Popovič *et al.*, 2014; Wamba *et al.*, 2017), gaining competitiveness (Côrte-Real, Oliveira and Ruivo, 2017), advancing business (Grover *et al.*, 2018), making better and faster decisions (Loebbecke and Picot, 2015; Ghasemaghaei, Ebrahimi and Hassanein, 2018), deriving functional value from business process improvement, product and service innovation, or customer experience and market enhancement (Grover *et al.*, 2018). In this sense, data science work's contribution to strategic value can be seen through the strategy-as-practice lens (Jarzabkowski *et al.*, 2016), with focus on the more minute mechanisms in data scientists' work and how they are intervowen with stategizing.

However, literature on the link between analytics outputs derived through data science work and strategic value is for now limited (Grover *et al.*, 2018; Koch, Chipidza and Kayworth, 2021). In other words, we lack a firm understanding of how data scientists ensure that the outputs of their work can contribute to valuable strategic outcomes, in the light of studies that show data can to lead to strategic failure instead (Aversa, Cabantous and Haefliger, 2018; Joshi *et al.*, 2021). The few studies that do empirically address this question offer important insights, for example Božič and Dimovski (2019) present a model linking the use of new insights derived from data science, innovation ambidexterity and firm performance. Koch and colleagues (2021) analyze how business managers make sense of new knowledge which leads to structural consequences for organizations. None of the extant studies though focus on how analytics outputs are created by data scientists–inside a data science team–in a way that aligns them with strategic value in the first place.

2.2 Data scientists at work

Extant literature offers insights into how data scientists generate analytics as part of their work, drawing from a range of skills, usually from the areas of programming languages, statistics, and domain knowledge (Donoho, 2017; Avnoon, 2021; Vaast and Pinsonneault, 2021). When conducting data preparation, analysis, and modelling, data scientists engage in exploratory programming (Hill *et al.*, 2016; Kery, Horvath and Myers, 2017), which is key under the circumstances of flexibility, discovery, and innovation, and its defining characteristic is "the practice of designing the goal *at the same time* as experimenting in code" (Kery and Myers, 2017, p. 25). Data scientists write code to prototype or experiment with different statistical approaches to obtain new knowledge from data, but they do not engineer working code to match a pre-defined specification–instead, the goal is open-ended and evolves together with the evolution of the code (Kery and Myers, 2017). This kind of work is highly iterative, exploratory and non-linear (Patel *et al.*, 2008), and relies heavily on professional judgment and decisions made by data scientists (Rule, Tabard and Hollan, 2018). As data scientists engage in obtaining, cleaning, profiling, analyzing and interpreting data (Rule, Tabard and Hollan, 2018), they may be tasked with data merging and cleaning, sampling, feature selection, defining metrics, building predictive models, defining ground truth, hypothesis testing, and even applying new knowledge or models to business (Kim *et al.*, 2016). At every step, they "try different versions of the same analysis, slowly improve analytical methods, and hit numerous 'dead ends' before finding an explanation that 'fits' the data" which "can make it difficult to perform an 'objective' analysis" (Rule, Tabard and Hollan, 2018, p. 2), while data scientists must be clear and transparent in their reasoning "if others are to understand, and ultimately trust their work" (Rule, Tabard and Hollan, 2018, p. 1).

This research into how data scientists derive new knowledge from data offers important insights, but does not necessarily place it against an organizational backdrop. Nascent IS studies attempt to fill this gap by investigating what happens at the intersection of data science work and organizations. For example, Ghasemaghaei and colleagues (Ghasemaghaei, Ebrahimi and Hassanein, 2018; Ghasemaghaei and Turel, 2021) discuss the practices of knowledge hiding that data scientists engage in, thus impacting how data science outputs are used in organizations. Joshi (2020) investigates how the practices of data scientists in the banking industry rely on both subjectivity and objectivity in the production of information. In an ethnography of how a machine learning-based analytics system was developed, van den Broek and colleagues (2021) show how knowledge produced within such systems relies on a hybrid practice combining machine learning and domain expertise in data science work. Parmiggiani and colleagues (2022) focus more specifically on finding and preparing the data that data scientists engage in to make data available for further analysis. This indicates that data scientists are actively involved in shaping the data, information, and new knowledge that subsequently circulates in the organization (Pachidi *et al.*, 2021). However, we still have a limited understanding of what data scientists do to ensure that the outputs that they contribute to organizations are linked with strategic value. To this end we ask: *How do data scientists create analytics outputs that contribute to strategic value?*

2.3 Craft as a lens to study data science work

Data scientists have strong feelings about and attachment to their emergent professional identity (Vaast and Pinsonneault, 2021), engaging in identity work that bridges scientist and engineer identities, draws from multiple theoretical orientations, foregrounds self-learning, and emphasizes domain knowledge (Avnoon, 2021). Data scientists have been found to seek positions of authority and status, and cultivate a global occupational community (Avnoon, 2021). The literature on how they approach their work summarized above is indicative of exploration underpinned by a masterful command of a range of skills needed to take care of the whole process from data collection to making new knowledge available, e.g., through visualization. These findings point towards craft, that is "a humanist approach to work that prioritizes human engagement over machine control" (Kroezen *et al.*, 2021, p. 502), as a plausible lens to study how data scientists organize their work to align it with strategic value. This approach embraces human engagement in work, and foregrounds individual judgment and involvement that cannot be replicated with a mechanical approach, or replaced by machines (Kroezen *et al.*, 2021; Raisch and Krakowski, 2021). Approaching work as craft grants workers control and autonomy over all facets of the work process (Hodson, 2010; Kroezen *et al.*, 2021), and thus the quality of work depends on the judgment, dexterity, and care of those who do it (Pye, 1995; Kroezen *et al.*, 2021). The perspective of craft, earlier applied to study, amongst others, software programming (Barley, 1996; Adler, 2015), entails a distinct approach to work that is based on specific skills and attitudes that differ it from other, more mechanical approaches. Craft skills encompass the mastery of technique, whereby individuals have exceptional competency over their work, all-roundedness that allows them full control over the entire process, and embodied expertise that is tacit and contextual (Barley, 1996; Kroezen *et al.*, 2021). Craft attitudes emphasize dedication, a full commitment to work and engagement in it for its own sake (Sennet, 2008), communality, that is attention to a shared occupational identity and purpose (Anteby, 2008), and exploration which requires tinkering and engaging with the complexity and ambiguity of tasks (Sennet, 2008; Kroezen *et al.*, 2021). We use the lens of craft together with its skills and attitudes as a sensitizing framework guiding our research into how data scientists align the new knowledge they create with organizational strategic value.

3 Research design

3.1 Research site

We conducted a single-case study (Yin, 2018) focusing on a data science team embedded in a people analytics department supporting the human resources (HR) function at a large, multinational technology company TeX (pseudonym). The team consists of 12 team members (6 data scientists and 6 research scientists of various levels of seniority) and one manager, and is nested in a department of around 40 data analysts, consultants, and business analysts working in people analytics supporting a 3,000-strong

HR function serving over 100,000 TeX employees around the world. The team is located within HR and as such is organized as an embedded data science team, working solely on HR-related data science projects and falling under HR in the organizational structure. Example projects include modeling attrition pressures using social graph analysis combined with other sources, or measuring worldwide employee satisfaction through surveys and natural language processing, both of which contribute to the strategic goal of HR to retain employees. In general, data scientists on the team are charged with analyzing data, building models, and developing data science products, such as dashboards and web applications, while research scientists tend to focus more on identifying suitable research questions, designing data collection, and overseeing the conduct of data science projects from a scientific standpoint. Although the team had existed within the department for a number of years, at the time of the study it was undergoing a period of structural and managerial change to better align it with the strategic goals of the organization. The initial team consisting of six data scientists was merged with a team of six research scientists and the combined teams were brought under the leadership of the Head of HR Data Science who previously managed the research science team. With a change in structure and leadership, the team identified an opportunity to shape its identity and align its work better with the strategic direction and goals of the HR function. We conducted our research during this period which gave us an unprecedented insight and access to informants who were evaluating their previous practices and re-designing their ways of working to ensure as high a strategic impact as possible. As such, the case is then revelatory in nature, and allowed us access to a contemporary phenomenon in depth and in its real-world context (Yin, 2018).

3.2 Data collection and analysis

Due to the interest of the project, we sought out a case that would give us access to a whole data science team that prioritized the strategic impact of its work. After obtaining ethical approvals, we gained access to the team through its manager, who introduced us to those team members who were interested in and available to participate, yielding a high response rate. Over a period of three months in mid-2021, we conducted 19 semi-structured interviews with 10 out of 13 employees, with 6 of them interviewed more than once. The interviews were conducted in a video format, with only voice recordings which were later transcribed. All interviews took place in a work-from-home setting which has traditionally been the mode of operating for the team due to its distributed nature, with team members across the USA and in Europe. In total, we collected around 1,160 minutes of interviews, broken down per position as shown in Table 1 below.

Position	**Main duties and background, interviewee ID**	**Number**	**Total interview time in minutes**
Head of HR Data Science	Managing the data science team, PhD in organizational psychology, Interviewee 1	1	157
Lead Data Scientist	Organizing work among data scientists, degree in statistics, Interviewee 2	1	177
Data Scientist	Working on analytics projects, degrees in psychology, statistics, PhD in operations research, programming and database management skills, Interviewee 3, 6, 7	3	392
Data Science Intern	Supporting analytics projects, degree in computer science, focusing on UX/UI, Interviewee 4	1	55
Research Scientist	Designing and conducting people analytics projects, PhD in econometrics, industrial psychology, MBA, Interviewee 5, 8, 9, 10	4	381

Table 1. Interview data collected.

In the interviews, we asked questions related to the skills and attitudes that informants needed to do their job together with their background, the way their work on various data science projects progressed, and how they ensured the value of their work as a team. Interviews with team members were an appropriate method for data collection as they enabled us to ask about team members' practices and changes they introduced to them to ensure the strategic value of their work. Interviewing enabled us to ask team members directly about their motivations and reasoning behind specific practices, and ask directed questions to elicit suitable information. As a result, we collected data explanatory in nature, yielding insights about participants' perspectives (Robson, 2011; Yin, 2018).

We analyzed the data using grounded theory methods (Glaser and Strauss, 1967; Charmaz, 2006; Urquhart and Fernandez, 2006; Wiesche *et al.*, 2017). We applied grounded theory methods in two forms, first to develop a model capturing data scientists' craft (by coding for craft skills such as dedication or communality), and second to theorize the process of data crafting at work. Because of the inductive nature of our study, we oscillated between data analysis and theory construction (Strauss and Corbin, 1990). We proceeded through multiple readings of the data collected as we analyzed and synthesized the findings (Eisenhardt, 1989; Beane and Orlikowski, 2015). As we did so, we relied on constant comparison, that is the process of constantly comparing units of data in constructs we identified with each other, and we engaged in memoing to capture our ideas of concepts, categories, and the relationships among them (Glaser, 1978; Wiesche *et al.*, 2017). We present an excerpt of our coding in Figure 1, and we report on findings below.

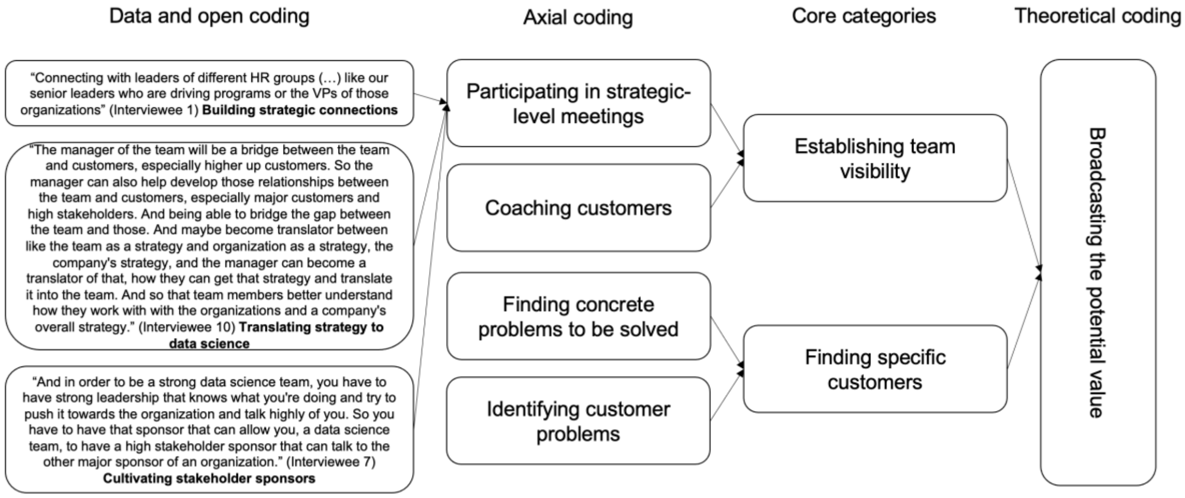

Figure 1. Excerpt of the coding process.

4 Findings: Doing strategically valuable data science

Our data is replete with evidence of craft as a suitable lens to capture the approach to ensuring the strategic value of data science work that the team was engaging in, that is craft as an approach to work that prioritizes human engagement over machine control (Kroezen *et al.*, 2021). We registered in the empirical data the presence of typically human craft skills (mastery of technique, all-roundedness, embodied expertise) and attitudes (dedication, communality, exploration). Data scientists saw themselves as craftspeople: "*as cliché as that sounds, I'm not in the camp of people calling us poets, but you have to be able to create something with a start, a middle, and an end*" (Interviewee 2).

The data science team at TeX was aware that their placement in the organization as an embedded data science team in a support function – human resources – entailed increased pressures on the value of work produced. As one interviewee expressed, for the organization as a whole if "*you sit in a support function that costs money [you need to] show something so we can make our business case for spending* *our money*", and for this reason "*in a company where you're the support function, and you're a data scientist, a good team understands business impact and business value*" (Interviewee 5). Leveraging the opportunities stemming from recent reorganization and new management, the team worked towards ensuring that the strategic value of their work would be evident for the function and the entire organization. We discuss how the team did it next.

4.1 Broadcasting the potential value of data science

One of the mechanisms of ensuring strategic value that the team deployed was broadcasting the potential value of their work outside of the team and towards the wider HR function. The need to engage this mechanism had been identified in the past by previous team managers and various more senior team members, and the team expressed the continuing need to promote, market, and sell the team to HR and beyond. Broadcasting potential value was mostly conducted through two sets of practices, establishing visibility, and finding customers.

Establishing team visibility was actively done by the Head, Lead, and a number of Research Scientists and past managers, and strongly advocated for by team members. The manager's role was identified as "*a translator between the team and the company's strategy*" (Interviewee 10), and as such the team emphasized that the strategic value of the work that it was able to deliver "*was something that had to be invented*" (Interviewee 3) drawing from business acumen, marketing, and vision skills, and then communicated directly to the wider function. To do so, more senior team members would be actively involved in "*connecting with leaders of different HR groups (…) like our senior leaders who are driving programs or the VPs of those organizations*" (Interviewee 1), and participate in higher-level meetings to "*help develop those relationships between the team and customers, especially major customers and high stakeholders*" (Interviewee 10). Participating in strategic-level meetings was an ongoing process, and a large portion of the manager's and more senior team members' time was dedicated to setting up meetings with high-level stakeholders to communicate the potential value of what the team could provide to them. To do so, all team members involved had to exhibit high levels of business acumen, as "*not understanding the business [is] the kiss of death*" (Interviewee 5).

Participating in meetings and establishing crucial connections fostering the visibility of the team often involved or led to the need to coach customers. In many instances, team members were able to garner interest in the possibilities that data science offered, but then realized that customers in the function were not aware of what data scientists did and what the team could provide: "*a lot of them [customers] can't visualize or don't understand what's possible*" (Interviewee 4) and "*people may be hesitant about what we are delivering for them, they maybe don't understand AI or data science and data science work*" (Interviewee 7). In some instances, team members would observe and shadow stakeholders and their HR teams, and hold regular meetings to understand the processes and show how they could contribute. Frequently, the manager and more senior team members were involved in showing past solutions, demonstrating previous projects, and explaining what the team could deliver for a particular group of customers.

More senior members of the team were also involved in establishing team visibility by proposing specific projects drawing from their own research, professional expertise and business acumen. This approach was rarer, but nonetheless sometimes deployed. As the manager explained, "*our team will sometimes propose our own research ideas based on what we're seeing in the sensing sources that we get data from*" (Interviewee 1) and "*those science projects should be with a customer in mind, even if you haven't talked to them yet*" (Interviewee 2). This happens when "*you know that it's time to not just think like a data scientist, but to think like and AI strategist*" so that it's possible to "*lay out a business case as to why you should use machine learning to tackle this problem, where the business case communicates the goals, the outcomes, the deliverables, and how you can use these deliverables for action planning*" (Interviewee 8). These research-driven, specific projects proposed were contributing to the overall visibility of the team, and required a large degree of exploration with respect to data, the organization, and potential projects that could be set up. These were essential to ensure that the way strategic value in HR was generated incorporated the work of the data science team.

While establishing team visibility was more dispersed and generally directed towards increasing the overall awareness, the team also engaged in more focused actions of **finding specific customers**. The interviewees emphasized that in the past, the team would sometimes create products first, and look for customers later, which would lead to "*failures or struggles because of creating products that no one's using*" (Interviewee 5). To avert this, the new approach was adopted of "*reaching out to people outside, to get them to buy into it before you even start building it*" (Interviewee 5). To do this, the team shifted its approach towards looking for specific customers and understanding their needs. Identifying customer problems to solve became an important part of the work done by the manager and more senior team members.

A significant amount of time, both of the manager and more senior team members, was spent finding out concrete problems that could be solved, which was equated with creating opportunities for doing valuable work, and then using their all-roundedness in business, technical, and domain knowledge to create projects out of problems with the leaders' buy-in. Identifying business problems to solve was often seen as a manger's role involving "*being a visionary, being able to see possibility in translating problems into possibilities, and looking at the different problems that our company has, and putting them together to see if you can create a solution*" (Interviewee 8). This created the need to understand the business and be attuned to customer needs. To be able to do this, team members had to have a certain level of knowledge in all aspects of data science work. Finding specific customers, especially among strategic stakeholders, fostered the alignment of the team's work with strategic value.

4.2 Cultivating a shared vision of value

Ensuring that the data science team would deliver strategically-valuable outcomes relied also on the mechanism of cultivating a shared vision of value within the team, that is ensuring that the promise of data science value that was broadcast to the organization was shared and bought into by the team. Ensuring agreement and unity within the team was seen as crucial because "*you're already fighting the environment you're trying to support, so you can't and don't want to have fighting within the team*" (Interviewee 3). Cultivating this shared vision going forward required the team to constantly build communality, align with strategy, and engage in continuing development.

Building communality was an important way in which the team socialized the shared vision. At a foundational level, this approach was built into how the team worked, by moving away from predominantly independent, individual contributors in the past towards relying on each other's skills to work "*in a mode that is essentially like everybody building components towards a shared or common vision*" (Interviewee 3). This distribution of expertise and diversity of backgrounds did not resonate with some previous managers who tended towards more focused approaches to skill building, but this resulted in frictions and silos within the team. In renewed attempts, new management moved towards the communal vision of work, resulting in projects where "*if [X] built some machine learning predictive component, and [Y] built some user interface, I now have to get it into production for them so we're all going to meet together, and I'm going to help [Y] and [X] stitch their parts together (…) so there's that, let's bring the pieces together*" (Interviewee 2). This entailed the fact that all team members were differently specialized and they could be drawn into specific pieces of work, largely dependent on their professional expertise, interests, and availability. The team became more closely knit together as a result, and various team members emphasized the need and possibility to complement each other, underpinned by mastery in their respective areas of expertise. Admitting that it was not possible for one person to have all skills at the needed level of mastery, the team communally distributed skills to be able to deliver strategically valuable products. At the same time, towards outside customers, the team as a whole offered a unified vision of how they were delivering that value because all team members were knowledgeable and involved enough in each other's work to speak about whole projects.

Supporting communality were overt and open arrangements of mentorship, both when socializing new members into the team, and when existing team members wanted to develop specific skills. In some instances, mentoring took the form of pair programming, or simply giving feedback and suggestions on each other's work: "*and then [X] is just now learning programming, and a lot of the time she'll ask me for advice, or we'll have little coding sessions together, and I'll try to help her get unstuck out of her problems*" (Interviewee 4), as an example of a more established team member learning a specific skill. This approach was well received within the team.

Communality was also built by establishing networks of expertise that allowed outside knowledge to enter the team. At the foundational level, networks were fostered within the team through weekly meetings where work was being shared, but an important step was to establish a weekly meeting for everyone in the HR function interested in data science that allowed the team to learn what other pockets of data scientists were working on and draw from their experience. Various team members built their own networks of experts whose knowledge they were learning from and contributing with to the team: "*the other thing you need to do is to establish a network inside and ideally outside of the company, you need to be tuned into what's going on*" (Interviewee 2). This was achieved through networking and attending industry events, and in the case of one team member, creating meetups and participating in a programming community in their country of residence. Building communality ensured that the work that the team was doing had the capacity and the right level to contribute to strategic value.

Aligning the team with the function's strategy was a constant, ongoing action that all team members were aware of. In order to ensure this constant alignment, the new manager started attempts to refine the operating model of the team. In the past, projects would typically find their way to the team through a variety of sources and requests, but this resulted in the team's work diverting from HR strategy. To counter this, the manager began to establish a new process whereby "*most of the work that we're doing should be aligned to HR strategic pillars, so it should be coming to my manager and myself through strategy leads*" (Interviewee 1) and then undergo prioritization and scheduling. Alignment also depended on ensuring that there was an agreement between what was strategically needed and what the team was able to deliver, and required some changes to the team's operating model, as well as upskilling and recruiting team members with specific skills, such as user experience or interface design. Ensuring alignment also required drawing boundaries around the work that the team was taking on. The team engaged in constant discussions around what constituted data science work that they should be taking on. In general, a number of team members expressed stronger views that their time should be reserved and protected for actual data science work, rather than supporting or less related tasks, as this was ensuring that the work was strategically valuable and it maintained the shared vision the team had built.

Engaging in continuing development was another way in which the team maintained the shared vision of how they were contributing to strategically valuable outcomes. Leveraging the networks of outside experts built, team members were able to keep track of the most valuable skills and then continually upskill themselves: "*I have a subscription to Medium for example, and I really like to read a few articles every day (...). So I think learning is a central part, I'm trying to even take some TeX courses that they publish every quarter*" (Interviewee 7). Team members recognized that there were some skills and areas of knowledge that were essential in their work that could have only been learned on the job, for example gaining subject matter expertise in HR and organizational psychology, learning the inner workings of the organization, or building up business acumen. The latter was especially desired and sought after.

Continuing development for data scientists also meant engaging in independent research projects that were not directly related to their work tasks: "*we try to preserve people's time for research, as well as not just being on projects (...) and we try to make sure everybody is growing their skill sets*" (Interviewee 2). It was an ambition for the team to reserve around 30% of time for independent research which could be stimulated by questions from customers, reading a paper, or coming up with completely novel ideas. Continuing development kept the team up-to-date with skills to enable them to contribute to value.

4.3 Creating value-adding data products

The team shared a view that data themselves were not intrinsically valuable for the function and the organization, and saw their role primarily as creating valuable tools, solutions, presentations, dashboards, and others, on the basis of data, that are aligned with the business needs of the organization. This view contrasted with some other experiences on data teams in other organizations. To create data products with impact, the team engaged in identifying the scope of data, productizing data, and co-designing the products with customers, which relied on bi-directional communication from the team to customers, and from customers to the team.

Identifying the scope of data was frequently deployed by the team and concerned not only acknowledging the technical limitations, but more importantly asserting and explaining to customers what the data were not and what they were not covering. This was seen as especially important within the HR function, where interviewees admitted that customers were quick to draw conclusions about what the data and analytics were saying without considering alternative explanations, for example around employee attrition, or sometimes not seeing that using some data could be against legal and ethical provisions. In these cases, data scientists felt it was their job to point out such issues to customers. There were also areas where data did not exist, so team members would need to think of ways of using alternative data, or collecting the data that are needed. Importantly, data scientists had to be aware of what the data themselves represented and how they related to the organization. In general, the team was aware that the data themselves were inherently limited, and they needed to be turned into informative products to be valuable within the organization. To point out such limitations, the team required a level of mastery of technical skills, and a certain level of expertise in understanding the organization.

Productizing data was an approach that permeated the team, whereby the team saw their work as creating analytics tools that were being offered as ready-to-use products to customers: "*we want to think about the end product in a way that the customers can engage and work with something*" (Interviewee 7). This shift in thinking was prompted by a few failed projects in the past, for example a career progression tool that was powerful behind the scenes, but had a basic interface "*so when it was demoed, people couldn't see the capability [that they would if we] would have dropped a user interface or skin onto it*" (Interviewee 3). As a result, the team focused more on ensuring that the end product the customer was receiving was a user-friendly and applicable tool, and adjusted the products to customer needs, which could vary from receiving a slide deck with data visualization, a CSV file, or a fully-fledged web-based dashboard. Productizing data required team members to engage in constant creating, building, and developing, and they often mentioned creative thinking that this work involved. This approach resulted in team members having a strong sense of ownership over the products they were developing. Productizing data also entailed a reduction of complexity through hiding away complex models and statistics in simpler interfaces, and moderating the level of involvement that customers had to have in terms of understanding what was happening behind the scenes. Developing working products for customers was seen as a way of expressing professional judgment in hiding away the unnecessary detail from customers who did not need to know more than required for the tool to be useful. Productizing data was important to feed into strategy, as more abstract data products were less prominent in strategy discussions.

Co-designing products with customers was how the team worked bi-directionally with customers to ensure that the actual data products that were developed were as strategically valuable as possible. Customers were usually involved throughout the whole process and asked to participate in regular meetings. The team adopted an agile approach with relying on user stories, developing minimum viable products, and holding demos of products every few weeks. Involving customers was seen as essential "*because as a data scientist, we do not know more than they know, they know way more than we know, but to think we're experts on their business and not listen to them will kill a project no matter how great you think it is*" (Interviewee 5). For this reason, the process was collaborative and involved customers, often senior stakeholders, and this involvement fostered the strategic value of the work done by the data science team.

5 Discussion: Data crafting in generating strategic value from analytics

Both research and practice suggest that organizations, despite dedicating not insignificant resources to data science efforts, struggle with deriving strategic value from these initiatives. Our study aimed at investigating this problem by analyzing the work of a data science team in a period of change that spurred efforts to closely align the work with the strategic objectives of the HR function the team was embedded in. We found that data scientists in the team at TeX create analytics outputs that contribute to strategic value through three mechanisms underpinned by skills and attitudes characteristic of craft (Kroezen *et al.*, 2021), and made possible by creating suitable conditions for craft by the organization. We term this approach to generating strategic value from analytics *data crafting*. We summarize it in Figure 2, and explain it in detail below.

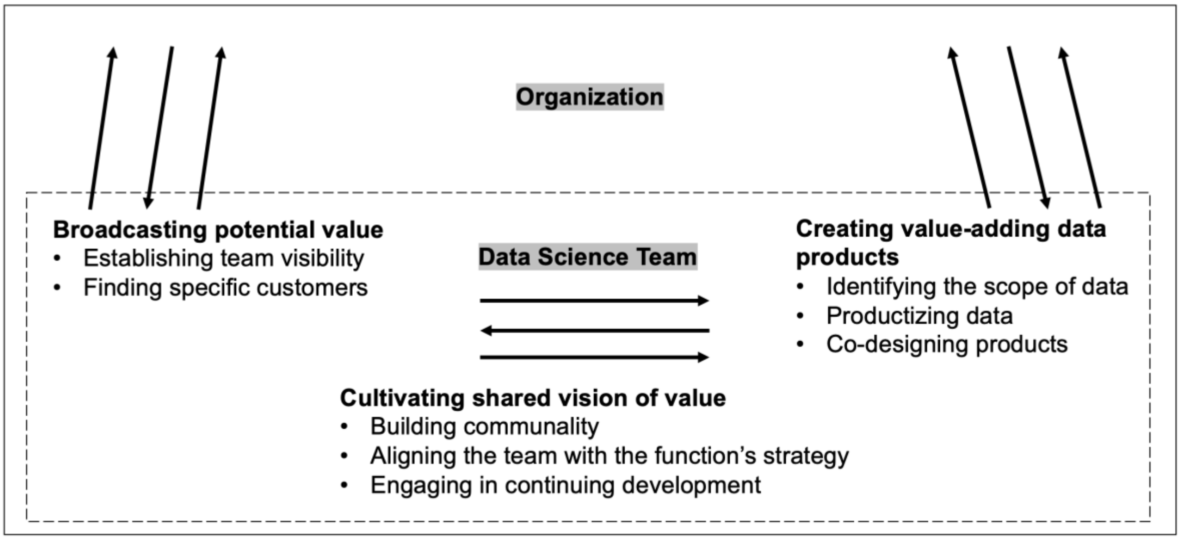

Figure 2. The mechanisms of data crafting.

Based on our findings, we define *data crafting* is an approach to creating strategically valuable analytics outputs by broadcasting the potential value of data science, cultivating a shared vision of value, and creating value-adding data products. The first mechanism focuses on promoting the potential value of data science in general towards potential customers in the organization, and it operates through establishing visibility within the function, and finding customers for data science work. Data scientists draw from craft skills here by showing the mastery of business acumen and knowledge obtained through years of experience in business roles, while exhibiting familiarity with the entire data science process. These are required to shape and communicate the value of data science to business customers. Data scientists who engage in broadcasting often bring their personalities, unique attitudes, and backgrounds into the work they do, and tend to share some of their occupational identity with customers, rather than seeing themselves purely as data scientists. They have an attitude of exploration and often make the first steps to be involved in strategic-level conversations and initiatives. To do so, the organization supports data scientists by allowing them to engage with high-level executives and senior management bi-directionally, that is by opening up channels through which the value can be broadcast and feedback received. At the same time, data scientists are given enough freedom to exercise professional judgment in assessing where and how data science may contribute the most to the organization's strategic goals.

The second mechanism operates by maintaining a shared vision of the value that the team provides by building communality, aligning the team's work with business strategy, and engaging in continuing development. Mastery of technique is required and expected of every data scientist, and they all have specific areas of specialization. While all data scientists know about all elements involved in data science work and are up to date with what other team members are working on, all-roundedness is distributed within the team, i.e., the data science team is overall in control over and involved in the entire making process (Kroezen *et al.*, 2021). Data scientists exhibit engagement in their work and identify with their professions even outside of their jobs. They cultivate a strong shared identity within the team and are engaged in fostering a community of practice, which is evident in the various meetings, events, and opportunities to share their work that they create. Exploration is an attitude to work that is fostered and rewarded within the team, with time protected for continuing development and independent research. To cultivate this shared vision, the team's cross-disciplinary character is accepted by the organization, and the variety of backgrounds of data scientists are not questioned, while team members are given the resources to engage in self-directed work.

The third mechanism relies on turning data into impactful products by identifying the scope of data, productizing data, and co-designing products with customers, and it requires bi-directional communication between the team and the customers. Mastery of technique concerns here a specific perspective on data science that prioritizes creating impactful data products, and for this all team members need to display awareness of the data product being created. For this reason, team members are actively involved in the entire process of creating the data product, and often display dedication and pride regarding actual products offered to customers. The team creates a communal environment with the customers to facilitate involving them in creating data products, and both data scientists and customers together explore the process of developing the product. This is supported by the organization allowing data scientists to draw on customers' time and resources to involve them in the co-creation process, and the team is trusted to propose and follow their preferred process.

5.1 Theoretical implications

The concept of data crafting allows us to revisit literature on deriving strategic value from data science that so far has viewed analytics and strategy as two separate components, with analytics outputs feeding into strategic value. However, we show that to deliver actionable insights (Zhang, Wang and Pauleen, 2017; Chen, Preston and Swink, 2021; Rana *et al.*, 2021) or new information (Hu, Kettinger and Poston, 2015; Someh, Shanks and Davern, 2019; Castillo *et al.*, 2021; Pan *et al.*, 2021) for strategic outcomes such as improved decision-making (Loebbecke and Picot, 2015; Ghasemaghaei, Ebrahimi and Hassanein, 2018) or performance (Popovič *et al.*, 2014; Wamba *et al.*, 2017), data science needs to be positioned within–not before or after–strategy.

The second condition for fostering data crafting, autonomy, ties into this as well. Much of literature assumes that data scientists work in the service of their customers in ways that define and constrain their ways of operating, outputs produced, and processes deployed. Our findings demonstrate that for analytics outputs to be strategically valuable, data science teams need a degree of autonomy to suggest potential contributions to value, engage in self-directed work, and propose approaches to creating data products. This is an alternative vision of a data scientist, rather than a lone, highly-specialized, individual contributor, towards a networked but autonomous professional. Data crafting requires these two conditions to be in place to enhance the likelihood of data science work contributing to strategic goals.

But data crafting also contributes to the understanding of the nature of the new forms of data that spurred interest in data science in organizations. The so-called big data have traditionally been seen as freer from human bias, the need for theory, or research design to collect them (Anderson, 2008; Davenport, Harris and Morrison, 2010), and as such they promised a more open access to data (Davenport and Harris, 2017). However, what we observe is that these kind of data often require even more human involvement from even more skilled and specialized professionals than previous, structured data that permeated organizations. Trained data scientists are needed to manage the volume, velocity, and variety of these data, but also to give them structure, purpose, and performative powers in organizations. Precisely because these data are unstructured, unpurposeful, hetegoreneous and performative (Kallinikos, Aaltonen and Marton, 2013; Constantiou and Kallinikos, 2015; Yoo, 2015; Grover *et al.*, 2020) they need to be carefully crafted into organizationally relevant data products by data scientists. Interestingly enough, data scientists themselves often relegate data to lower importance and emphasize that it is the crafting of a meaningful, impactful data product that can make a contribution towards strategic value. In other words, it is the craft that happens around data that leads to strategic value.

5.2 Managerial implications

Organizations invest increasing volumes of resources into data science and analytics, but often struggle to derive strategic value from these initiatives. Our research offers managers a better understanding, through the three mechanisms we identified, of how such value can be ensured. We identified two craft conditions that organizations should create to foster the craft approach to work: positioning and autonomy. The right positioning of employees or teams and their autonomy within an organization can enable them to exercise craft skills and approach work with craft attitudes. The concept of data crafting that we propose captures the mode of work of data science teams to foster the right environment to make their work strategically relevant: data science based on digital technologies may require espousing a craft approach to ensure that value is derived.

6 Conclusions

In this paper, we offer a view from within a data science team to elucidate how analytics outputs are created to ensure their strategic value. We studied the work of data scientists who deploy three mechanisms of data crafting, an approach to data science work that supports creating strategically valuable analytics outputs. Our findings shed light on the pathway from data to strategic value, and contribute to the understanding of the role of data and data science in organizations. The limitations of our study, most notably the case context and the positioning of the team as an embedded data science team in a support function, open up avenues for further research. For example, a study focusing on a centralized data science team striving for the strategic value of its work can test and extend our model. Another complementary study could focus on organizational customers of analytics outputs to understand creating the strategic value of data science work from their perspective.

References

Abbasi, A., Sarker, S. and Chiang, R. H. L. (2016). "Big data research in information systems: Toward an inclusive research agenda." *Journal of the Association for Information Systems* 17(2), 1–32.

Adler, P. S. (2015). "Community and Innovation: From Tönnies to Marx". *Organization Studies* 36(4), 445–471.

Alaimo, C. and Kallinikos, J. (2017). "Computing the everyday: Social media as data platforms." *The Information Society* 33(4), 175–191.

Anderson, C. (2008). *The End of Theory: The Data Deluge Makes The Scientific Method Obsolete*, *Wired*.

Anteby, M. (2008). "Identity incentives as an engaging form of control: Revisiting leniencies in an aeronautic plant." *Organization Science* 19(2), 202–220.

Aversa, P., Cabantous, L. and Haefliger, S. (2018). "When decision support systems fail: Insights for strategic information systems from Formula 1." *Journal of Strategic Information Systems* 27(3), 221–236.

Avnoon, N. (2021). "Data Scientists' Identity Work: Omnivorous Symbolic Boundaries in Skills Acquisition." *Work, Employment and Society* 35(2), 332–349.

Baesens, B. *et al.* (2016). "Transformational Issues of Big Data and Analytics in Networked Business." *MIS Quarterly* 40(4), 807–818.

Barley, S. R. (1996). "Technicians in the Workplace: Ethnographic Evidence for Bringing Work into Organizational Studies." *Administrative Science Quarterly* 41(3), 404–441.

Beane, M. and Orlikowski, W. J. (2015). "What difference does a robot make? The material enactment of distributed coordination." *Organization Science* 26(6), 1553–1573.

Berente, N., Seidel, S. and Safadi, H. (2019). "Data-driven computationally intensive theory development." *Information Systems Research* 30(1), 50–64.
Bhimani, A. (2015). "Exploring big data's strategic consequences." *Journal of Information Technology* 30(1), 66–69.
Božič, K. and Dimovski, V. (2019). "Business intelligence and analytics use, innovation ambidexterity, and firm performance: A dynamic capabilities perspective." *Journal of Strategic Information Systems* 28(4), 101578.
van den Broek, E., Sergeeva, A. and Huysman, M. (2021). "When the machine meets the expert: An ethnography of developing AI for hiring." *MIS Quarterly* 45(3), 1557–1580.
Castillo, A. *et al.* (2021). "Impact of social media on the firm's knowledge exploration and knowledge exploitation: The role of business analytics talent." *Journal of the Association for Information Systems* 22(5), 1472–1508.
Charmaz, K. (2006). *Constructing Grounded Theory*. Thousand Oaks, CA: SAGE Publications.
Chen, D. Q., Preston, D. S. and Swink, M. (2015). "How the use of big data analytics affects value creation in supply chain management." *Journal of Management Information Systems* 32(4), 4–39.
Chen, D. Q., Preston, D. S. and Swink, M. (2021). "How big data analytics affects supply chain decision-making: An empirical analysis." *Journal of the Association for Information Systems* 22(5), 1224–1244.
Chen, H., Chiang, R. H. L. and Storey, V. C. (2012). "Business Intelligence and Analytics: From Big Data to Big Impact." *MIS Quarterly* 36(4), 1165–1188.
Chiang, R. H. L. *et al.* (2018). "Special Issue: Strategic Value of Big Data and Business Analytics." *Journal of Management Information Systems* 35(2), 383–387.
Constantiou, I. D. and Kallinikos, J. (2015). "New games, new rules: Big data and the changing context of strategy." *Journal of Information Technology* 30(1), 44–57.
Côrte-Real, N., Oliveira, T. and Ruivo, P. (2017). "Assessing business value of Big Data Analytics in European firms." *Journal of Business Research* 70, 379–390.
Cybulski, J. L. and Scheepers, R. (2021). "Data science in organizations: Conceptualizing its breakthroughs and blind spots." *Journal of Information Technology* 36(2), 154–175.
Davenport, T., Barth, P. and Bean, R. (2012). "How "big data" is different." *MIT Sloan Management Review* 54, 43–46.
Davenport, T. H. (2018). "From analytics to artificial intelligence." *Journal of Business Analytics* 1(2), 73–80.
Davenport, T. H. and Harris, J. (2017). *Competing on analytics: Updated, with a new introduction: The new science of winning*. Harvard Business Press.
Davenport, T. H., Harris, J. and Morrison, R. (2010). *Analytics at Work: Smarter Decisions, Better Results*. Harvard Business School Press.
Donoho, D. (2017). "50 Years of Data Science." *Journal of Computational and Graphical Statistics* 26(4), 745–766.
Eisenhardt, K. M. (1989). "Building Theories from Case Study Research." *The Academy of Management Review* 14(4), 532–550.
Ghasemaghaei, M., Ebrahimi, S. and Hassanein, K. (2018). "Data analytics competency for improving firm decision making performance." *Journal of Strategic Information Systems* 27(1), 101–113.
Ghasemaghaei, M. and Turel, O. (2021). "Possible negative effects of big data on decision quality in firms: The role of knowledge hiding behaviours." *Information Systems Journal* 31(2), 268–293.
Glaser, B. (1978). *Theoretical Sensitivity: Advances in the Methodology of Grounded Theory*. Mill Valley, CA: Sociology Press.
Glaser, B. G. and Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine.
Grover, V. *et al.* (2018). "Creating Strategic Business Value from Big Data Analytics: A Research Framework." *Journal of Management Information Systems* 35(2), 388–423.
Grover, V. *et al.* (2020). "The perils and promises of big data research in information systems." *Journal of the Association for Information Systems* 21(2), 268–291.
Hill, C. *et al.* (2016). "Trials and tribulations of developers of intelligent systems: A field study."

Proceedings of IEEE Symposium on Visual Languages and Human-Centric Computing, VL/HCC, 2016-Novem, 162–170.

Hodson, R. (2010). "Work group effort and rewards: The roles of organizational and social power as context." *Organization Studies* 31(7), 895–916.

Hu, T., Kettinger, W. J. and Poston, R. S. (2015). "The effect of online social value on satisfaction and continued use of social media." *European Journal of Information Systems* 24(4), 391–410.

Jarzabkowski, P. *et al.* (2016). "On the risk of studying practices in isolation: Linking what, who, and how in strategy research." *Strategic Organization* 14(3), 248–259.

Joshi, M. *et al.* (2021). "Why So Many Data Science Projects Fail to Deliver." *MIT Sloan Management Review* Spring 2021.

Joshi, M. P. (2020). "Custodians of Rationality: Data Science Professionals and the Process of Information Production in Organizations." *AMCIS 2020 Proceedings*, 0–1.

Kallinikos, J., Aaltonen, A. and Marton, A. (2013). "The Ambivalent Ontology of Digital Artifacts." *MIS Quarterly* 37(2), 57–370.

Kery, M. B., Horvath, A. and Myers, B. (2017). "Variolite: Supporting exploratory programming by data scientists." *Conference on Human Factors in Computing Systems - Proceedings*, 2017-May, 1265–1276.

Kery, M. B. and Myers, B. A. (2017). "Exploring exploratory programming." *Proceedings of IEEE Symposium on Visual Languages and Human-Centric Computing, VL/HCC*, 2017-Octob, 25–29.

Kim, M. *et al.* (2016). "The emerging role of data scientists on software development teams." *Proceedings - International Conference on Software Engineering*, 14-22-May-, 96–107.

Koch, H., Chipidza, W. and Kayworth, T. R. (2021). "Realizing value from shadow analytics: A case study." *Journal of Strategic Information Systems* 30(2), 101668.

Kroezen, J. *et al.* (2021). "Configurations of craft: Alternative models for organizing work." *Academy of Management Annals* 15(2), 502–536.

Kuhn, K. M. and Galloway, T. L. (2015). "With a Little Help From My Competitors: Peer Networking Among Artisan Entrepreneurs." *Entrepreneurship: Theory and Practice* 39(3), 571–600.

Lehrer, C. *et al.* (2018). "How Big Data Analytics Enables Service Innovation: Materiality, Affordance, and the Individualization of Service." *Journal of Management Information Systems* 35(2), 424–460.

Leidner, D. E. and Tona, O. (2021). "The care theory of dignity amid personal data digitalization." *MIS Quarterly* 45(1), 343–370.

Loebbecke, C. and Picot, A. (2015). "Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda." *Journal of Strategic Information System* 24(3), 149–157.

Mikalef, P. and Krogstie, J. (2020). "Examining the interplay between big data analytics and contextual factors in driving process innovation capabilities." *European Journal of Information Systems* 29(3), 260–287.

Monteiro, E. and Parmiggiani, E. (2019). "Synthetic knowing: The politics of the Internet of Things." *MIS Quarterly* 43(1), 167–184.

Müller, O., Fay, M. and vom Brocke, J. (2018). "The Effect of Big Data and Analytics on Firm Performance: An Econometric Analysis Considering Industry Characteristics." *Journal of Management Information Systems* 35(2), 488–509.

Pachidi, S. *et al.* (2021). "Make Way for the Algorithms: Symbolic Actions and Change in a Regime of Knowing." *Organization Science* 32(1), 18–41.

Pan, S. L. *et al.* (2021). "Sustainability Design Principles for a Wildlife Management Analytics System: An Action Design Research." *European Journal of Information Systems* 30(4), 452–473.

Parmiggiani, E., Østerlie, T. and Almklov, P. G. (2022). "In the Backrooms of Data Science." *Journal of the Association for Information Systems* 23(1), 139–164.

Patel, K. *et al.* (2008). "Investigating statistical machine learning as a tool for software development." *Conference on Human Factors in Computing Systems - Proceedings*, 667–676.

Popovič, A. *et al.* (2014). "How information-sharing values influence the use of information systems: An investigation in the business intelligence systems context." *Journal of Strategic Information Systems* 23(4), 270–283.

Pye, D. (1995). *The nature and art of workmanship*. London: The Herbert Press.

Raisch, S. and Krakowski, S. (2021). "Artificial Intelligence and Management: The Automation-Augmentation Paradox." *Academy of Management Review* 46(1), 192–210.

Rana, N. P. *et al.* (2021). "Understanding dark side of artificial intelligence (AI) integrated business analytics: assessing firm’s operational inefficiency and competitiveness." *European Journal of Information Systems*, 1–24.

Ransbotham, S. *et al.* (2016). "Special section introduction: Ubiquitous IT and digital vulnerabilities." *Information Systems Research* 27(4), 834–847.

Robson, C. (2011). *Real world research: a resource for users of social research methods in applied settings*. Padstow: John Wiley.

Rule, A., Tabard, A. and Hollan, J. D. (2018). "Exploration and explanation in computational notebooks." *Conference on Human Factors in Computing Systems - Proceedings*, 2018-April, 1–12.

Saar-Tsechansky, M. (2015). "The business of business data science in IS journals." *MIS Quarterly:* 39(4), iii–vi.

Sennet, R. (2008). *The craftsman*. Yale University Press.

Sharma, R., Mithas, S. and Kankanhalli, A. (2014). "Transforming decision-making processes: A research agenda for understanding the impact of business analytics on organisations." *European Journal of Information Systems* 23(4), 433–441.

Someh, I., Shanks, G. and Davern, M. (2019). "Reconceptualizing synergy to explain the value of business analytics systems." *Journal of Information Technology* 34(4), 371–391.

Strauss, A. and Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Newbury Park, CA: Sage.

Teo, T. S. H., Nishant, R. and Koh, P. B. L. (2016). "Do shareholders favor business analytics announcements?" *Journal of Strategic Information Systems* 25(4), 259–276.

Urquhart, C. and Fernandez, W. (2006). "Grounded Theory Method: The Researcher as Blank Slate and Other Myths." in *Proceedings of the 27th International Conference on Information Systems.* Milwaukee, WI.

Vaast, E. and Pinsonneault, A. (2021). "When Digital Technologies Enable and Threaten Occupational Identity: the Delicate Balancing Act of Data Scientists." *MIS Quarterly* 45(3), 1087–1112.

Wamba, S. F. *et al.* (2017). "Big data analytics and firm performance: Effects of dynamic capabilities." *Journal of Business Research* 70, 356–365.

Wiesche, M. *et al.* (2017). "Grounded Theory Methodology in Information Systems Research." *MIS Quarterly* 41(3), 685–701.

Woerner, S. L. and Wixom, B. H. (2015). "Big data: Extending the business strategy toolbox." *Journal of Information Technology* 30(1), 60–62.

Yin, R. K. (2018). *Case Study Research and Applications. Design and Methods*. Thousand Oaks, CA: SAGE Publications.

Yoo, Y. (2015). "It is not about size: A further thought on big data." *Journal of Information Technology* 30(1), 63–65.

Zhang, T., Wang, W. Y. C. and Pauleen, D. J. (2017). "Big data investments in knowledge and non-knowledge intensive firms: what the market tells us." *Journal of Knowledge Management*. 21(3), 623–639.