BFJ
126,13

Received 5 August 2023
Revised 8 March 2024
Accepted 3 April 2024

Seeding innovation: the role of internal and external digital data in agri-food product innovation

Moreno Frau and Tamara Keszey
Department of Marketing Management,
Institute of Marketing and Communication Sciences,
Corvinus University of Budapest, Budapest, Hungary

Abstract

Purpose – Since previous literature provides fragmented and conflicting results about the use of digital data for product innovation, the article aims to comprehensively explore and shed light on how agri-food firms utilise external and internal digital data sources when dealing with different product innovations, such as incremental, architecture and radical innovation.

Design/methodology/approach – This paper adopts an exploratory multiple-case study and a theory-building process, focussing on the agri-food industry. We collected primary and secondary data from eight manufacturing companies.

Findings – The findings of this research show an empirical framework of six agri-food firms' digital data utilisation behaviours: the supervisor, the passive supervisor, the developer, the passive developer, the pathfinder and the conjunction behaviour. These digital data utilisation behaviours vary according to a combination of data sources, such as internal data related to inside phenomenon measures (e.g. data generated by sensors installed in the production plan) or external data (e.g., market trends, overall sector sales), and innovation purposes.

Practical implications – This article offers guiding principles that assist agri-food companies when utilising internal and external digital data sources for specific product innovation outcomes such as incremental, architectural and radical innovation.

Originality/value – The significance of external and internal data sources in stimulating product innovation has garnered substantial attention within academic discussions, highlighting the critical importance of analysing digital data for driving such innovation. Nonetheless, the predominant approach is to study a single innovation outcome through the lens of digital technology. In contrast, our study stands out by adopting a fundamental perspective on data sources, enabling a more nuanced explanation of the overall product innovation outcomes within the agri-food sector.

Keywords Product innovation, Data source, Digital data, Digital transformation, Agri-food industry, Multiple case study

Paper type Research paper

1. Introduction

Foodie is a food processing company that has embarked on a digital transformation (DT) journey. Armed with tradition and innovation, they dived into DT through the Internet of Things (IoT), cloud computing, and their deep culinary expertise to meticulously craft recipes using internal data, aiming to innovate flavours to new heights radically. As a result, Foodie failed to meet customers' tastes by ignoring the data from the external environment, such as customers' trends.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 101030185. This article is an outcome of the project FooDization.

British Food Journal
Vol. 126 No. 13, 2024
pp. 286-302
Emerald Publishing Limited
0007-070X
DOI 10.1108/BFJ-08-2023-0687

On the other hand, exploring external data sources with predictive analytics revealed a world of opportunities that Foodie could not seize once they just wanted to innovate their products incrementally. Yet, the sheer complexity of the data sources left them adrift. Foodie behaved incoherently when employing data sources for different product innovation purposes. This short story-telling introduction is emblematic of an issue linked with the DT, which is radically changing the global economy (Matzler *et al.*, 2018) and the company's capabilities to deal with data for product innovation and overall performance (Ferraris *et al.*, 2019).

DT is the organisational change triggered and shaped by the widespread diffusion of digital technologies (Hanelt *et al.*, 2021). The agri-food industry is no exception, as "computers are now used in all agriculture-related processes, from machinery to decision-making systems, through the use of robots, sensors and cyber-physical systems technologies" (Konfo *et al.*, 2023, p. 1). The agri-food business faces a significant shift, primarily driven by digital technologies which enable real-time data for product innovation (Frau *et al.*, 2022b). In this manuscript, we use the term product innovation, including exploratory and exploitative product innovation such as new product development, radical innovation and incremental development of existing products (Danneels, 2002). Mainly, DT has generated the proliferation of digital data—directly caught in a digital form (Piccoli and Watson, 2008). Digital data can be internally related to inner phenomenon measures (e.g., generated by sensors installed in the production plant) or externally related to a company's operational environment, such as market trends.

Prior studies investigating the impact of digital technologies on food product innovation have considerably increased the knowledge of how agri-food companies employ digital data (e.g., Beckeman *et al.*, 2013; Varese and Cane, 2017; Frau *et al.*, 2022a). For instance, using precision technologies, which rely on digital data to innovate products, farmers can improve crop yields and reduce inputs such as water and fertiliser (Romanello and Veglio, 2022). Additionally, other studies highlight the importance of utilising machine learning algorithms to analyse extensive datasets and forecast potential product innovation issues in advance (Belaud *et al.*, 2019). Also, predictive analytics can enhance processing and production innovation, minimise wastage, and elevate food safety and quality standards (Oltra-Mestre *et al.*, 2021). Digital data collection can also drive the innovation of new technologies that can transform the agri-food industry. A recent trend is the adoption of drones and sensors for data gathering in the agri-food industry, potentially improving product management and innovativeness (Kör *et al.*, 2022). Similarly, digital data and analytics can optimise production and reduce environmental impacts in innovative farming systems (Musa and Basir, 2021). In conclusion, prior literature suggests that *studying digital data in the agri-food industry is essential for driving product innovation.*

Nevertheless, despite the significance of this topic and the efforts made by earlier studies, it is still unclear how the agri-food firms' data utilisation behaviour varies according to the data sources (internal vs external data) and product innovation outcomes, such as radical or incremental (Zambon *et al.*, 2019; Demartini *et al.*, 2018). Distinguishing between internal and external data is pivotal in product innovation because different data sources lead to different innovation outcomes (Keszey, 2018). For example, external data sources lead to accidental knowledge leaking and lower radical innovation performance. The same results are not found for incremental innovation, implying that it is especially detrimental to radical innovation (Ritala *et al.*, 2018). On the other hand, Maes and Sels (2014) found that radical innovation is positively affected by both internal and external sources. Finally, a study on incremental innovation suggests that firms with a broad knowledge base benefit more from the diversity of internal data sources. Conversely, companies with extensive depth of knowledge tend to get greater advantages from the heterogeneity of external data sources rather than internal ones when pursuing incremental innovations (Du, 2021). Despite the worthy prior research efforts, concrete answers to the change in agri-food behaviour towards product innovation outcomes remain somewhat ambiguous, leaving a considerable gap in the literature (Oltra-Mestre *et al.*, 2021).

The present study aims to close this gap by empirically investigating how internal and external data utilisation leads to different product innovation results amongst agri-food companies. Therefore, the present study answers the following research question: "How do agri-food companies employ various data sources to drive product innovation?"

To fill the gap in the literature, we follow a theory-building process and employ an exploratory multiple-case study as the research methodology focusing on food manufacturers (Eisenhardt and Graebner, 2007). The findings of this study offer an empirical map made of six behaviours put in place by agri-food firms when dealing with data source utilisation for product innovation: the supervisor, the passive supervisor, the developer, the passive developer, the pathfinder, and the conjunction behaviour.

Compared with the prior research's fragmented view, our research adds to the ongoing academic discussion regarding the influence of digital technologies on product innovation within the agri-food sector in three key aspects. Firstly, we introduce a fresh analytical viewpoint to the existing literature by delineating between external and internal digital data. Secondly, our study sheds light on the underlying patterns concerning how companies utilise and merge digital data from diverse sources. Lastly, our research contributes to a more detailed and holistic understanding of how various digital data sources lead to distinct forms of product innovation. In terms of management impact, the behavioural map detailing the utilisation of data sources for product innovation serves as a guiding tool. It offers agri-food companies direction in effectively employing internal and external data sources in alignment with their innovation objectives.

The rest of this article is structured as follows. First, in the theoretical background, we analyse literature about the effects of digitalisation and the use of related technologies in the agri-food sector for product innovation, focusing on digital data sources. Then, in the second section, we introduce the methodology employed to conduct the study. Afterwards, we systematically show the relevant findings. Lastly, in the discussions section, we explain the paper's key theoretical contributions, important implications for practitioners, research limitations and suggestions for future research.

2. Digital transformation in the agri-food product innovation: technologies and digital data sources

Digital transformation dramatically impacts the agri-food industry by providing a new generation of equipment known as digital technologies (Adamashvili *et al.*, 2020; Steenis and Fischer, 2016). Digital technologies encompass software applications, hardware devices, and communication networks that facilitate data generation, acquisition, storage, analysis, and exploitation (Frau *et al.*, 2022a; Konfo *et al.*, 2023).

Nowadays, several types of digital technology are regularly used in the agri-food industry; some of the most impactful are Artificial Intelligence (AI), the IoT, and Big Data (Büchi *et al.*, 2020). Such technologies have created an agile environment for product innovation (Frau *et al.*, 2022b). In agri-food product innovation, AI, which refers to the ability of machines to acquire knowledge and make informed decisions by processing data, plays a significant role. It can automate innovation tasks like forecasting market and production trends. AI aids in preserving biodiversity and increasing innovativeness (Lezoche *et al.*, 2020). IoT involves the integration of sensors into tangible objects, allowing them to connect to wireless networks using the Internet Protocol (Ben Ayed *et al.*, 2022). In agri-food product innovation, several IoT applications enhance product performance by monitoring parameters such as temperature, humidity, and colour (Endres *et al.*, 2022). With AI and IoT-equipped production lines, agri-food companies are increasingly becoming data-enabled and data-driven, with vast availability of data sources, which leads to Big Data (Adamashvili *et al.*, 2021; Ferraris *et al.*, 2019). Big Data, abbreviated as BD, denotes vast and complex data beyond the processing capabilities of traditional techniques (Hassoun *et al.*, 2022). BD is distinguished by five Vs: volume, velocity, variety, veracity, and value, crucial for enhancing firms' performance (Ferraris *et al.*, 2019). For example, integrating BD in the agri-food sector is significant for developing innovative services complementary to core products (Ben Ayed *et al.*, 2022). The new data-driven agri-food industry is adopting data analysis techniques to extract meaningful information, making the sector more innovative (Kamble *et al.*, 2020). Therefore, previous agri-food literature has importantly explained how product innovation outcomes highly depend on data.

Nevertheless, only a few researchers examined the innovation of agri-food products from the fundamental point-of-view of *the data sources*. The role of external and internal data sources in driving product innovation has been subject to considerable scrutiny in academic discourse. Keszey (2018) posits that these sources contribute to innovation in distinct ways. Whilst external data sources are often hailed for their potential to foster innovation, Ritala *et al.* (2018) caution against their indiscriminate use, highlighting potential knowledge leakage and reduced performance in radical innovation contexts, albeit with less impact on incremental innovation. However, conflicting findings emerge from other studies (e.g., Maes and Sels, 2014; Hutchinson, 2020; Du, 2021), indicating a more nuanced relationship. In export activities, leveraging external data sources, such as insights into customers' needs in foreign markets, emerges as a pivotal driver of new product innovation (Li and Tamer Cavusgil, 2000). Moreover, continuous data collection on customers, market trends, competitors, operational environment, and technological advancements is crucial for fostering employees' capacity to contribute effectively to product development (Laforet, 2009). This emphasis on data utilisation extends to B2B markets, where customer involvement in external data analysis facilitates innovation, particularly in new product development (Zhang and Xiao, 2020). However, Maes and Sels (2014) suggest that internal and external data sources positively influence radical innovation. This underscores the importance of combining diverse data sources, both internal and external and leveraging AI tools to generate actionable insights that drive incremental advancements and the creation of new products (Hutchinson, 2020). Notably, Du (2021) suggests that firms with a broad knowledge base benefit more from internal data diversity in pursuing incremental innovations, whilst those with extensive depth of knowledge derive greater advantages from the heterogeneity of external data sources.

Although there has been a significant increase in research papers on internal and external digital data sources, only a few are related to the agri-food industry (Zambon *et al.*, 2019). Likewise, Demartini *et al.* (2018) found few studies including the keywords 'food' and 'digital', with most focusing on employing single technologies for specific innovation aims. Some studies even excluded food and agriculture from their analysis due to their incomparability with other manufacturing sectors (Müller *et al.*, 2020). However, despite the potential of digital data sources to enhance producer innovation, the research landscape remains lacking in agri-food-related studies (Oltra-Mestre *et al.*, 2021). In conclusion, former research did not exhaustively examine how internal and external data sources can influence the product innovation process differently (Schweitzer *et al.*, 2019).

3. Methodology

The article delivers a theoretical framework regarding how agri-food firms behave when using digital data for product innovation. Because digital transformation in agri-food companies has received less academic attention, we opted for an exploratory multiple-case study design (Eisenhardt and Graebner, 2007). We based the theoretical development on the empirical data to provide a comprehensive reply to the research question.

3.1 Research sample and case selection

We employed the theoretical sampling approach "to choose cases which are likely to [. . .] extend the emergent theory" (Eisenhardt and Graebner, 2007, p. 537)", following these criteria: the presence of digital technologies and product innovation activities. Regarding digital technologies, we involved companies that have invested in and implemented digital infrastructure such as IoT devices, data analytics platforms, and digital supply chain management systems. We also consider the extent to which digital technologies are integrated into various aspects of the company's operations, including production, supply chain management, and marketing strategy (Frau *et al.*, 2020). Finally, we checked whether the company could collect and analyse digital data relevant to agri-food production and product innovation. Concerning product innovation, we prioritised companies with a history of launching new products or introducing considerable improvements to existing ones.

The multiple case-study design requires gathering and matching data from several cases. We selected cases from the agri-food because this industry is highly representative and informative of the analysed phenomenon (Yin, 2009) and is considered suitable for addressing the theoretical purposes and research question. Also consistent with the research objective of investigating the use of digital data to facilitate product innovation, we selected cases from food processing firms because the process of innovating products is more observable within manufacturing companies than in the agricultural or distribution sectors. Whilst product innovation does occur in the latter sectors, it is more challenging to observe. We focused on Hungarian and Italian agri-food companies because they belong to the group of moderate innovators according to the European Innovation Scoreboard 2023, which makes them comparable with other EU Mediterranean construes such as Spain, Portugal, Greece, Malta, and EU central eastern countries like Estonia, Slovenia, Czechia, Lithuania, (European Commission, 2023).

3.2 Data collection

We gathered data from primary and secondary sources. Primary data consisted of semi-structured interviews with key informants selected within the company because they are managers leading in the firm's digital transformation strategy, product innovation process, or data analysis (e.g. CEOs, IT, R&D and Digital transformation specialists). Regarding the secondary data, we had access to archival data such as technology and product innovation strategies, firm social media pages (e.g., Facebook, LinkedIn, Twitter and Instagram) and official websites (see Table 1).

We employed an interview protocol of twelve questions and nine sub-questions organised into three units. In the first section, we asked preliminary questions about the company, the interviewee's role in the organisation, and the research context. For example, we asked: Could you please introduce the company and its products and explain your role in the organisation? What machinery does your company use for food processing? Please describe the most significant technological change in the last ten years that forced your company to change accordingly.

In the second section, the questions focus on the firm's technologies to create digital data. For instance, we enquired: What kind of data did food processing machinery create before introducing the new technologies? Which analysis did the firm perform on food processing data before introducing the new technologies? How did the firm use the information generated by the food processing data analysis before introducing the new technologies?

The third and final section asks questions regarding how the company employed digital data to innovate its food products. For example, what is the company's process for product innovation? How does the information generated by food processing influence the decisions

Table 1. Overview of the case studies

Case study	Business area	Case description	Size	Country	Key informant (experience in years)	Interview (minutes)	Technology adopted
1	Fruits and vegetable processing	The firm processes bio and local fruits to produce pulps, smoothies, juices, and vegetable products such as tofu, tempeh, and seitan	Medium-size	Hungary	CEO (20)	72	Smart Packaging (RFID tags and QR codes)
2	Organic farm	The firm produces and commercialises citrus fruits, kiwis, jams, marmalade, and juices	Small	Italy	CEO (8)	35	Mobile applications and online shop
3	Olive oil production	A cooperative firm that produces various extra virgin olive oil	Large	Italy	IT Specialist (11)	38	IoT, cloud computing
4	Dairy production	The firm specialises in dairy production, agricultural, and cow breeding sectors. It integrates the main activities with the production of raw materials and transforms sewage into electricity	Medium-size	Hungary	R&D (15)	51	IoT, cloud Computing, and analytics
5	Dairy production	A cooperative firm of shepherds that transforms and distributes cow milk obtained from the members' farms	Medium-size	Italy	Directing Manager (6)	45	Supply chain management systems
6	Poultry products	The firm handles the whole integrated production cycle the raw materials selection, breeding units, vivarium, feeding plants, food processing, packaging, and commercialisation	Large	Italy	Digital transformation specialists (18)	59	IoT, AI, analytics, and supply chain management systems, QR codes
7	Fresh pasta	The firm produces fresh pasta for organised large-scale distribution	Small	Hungary	CEO (5)	37	Cloud computing, data analytics platforms
8	Dry pasta and rusks production	The firm specialises in producing numerous categories and shapes of dry pasta and biscuits	Small	Italy	Production Manager (9)	55	IoT, analytics, and production management systems

Source(s): Authors' own elaboration

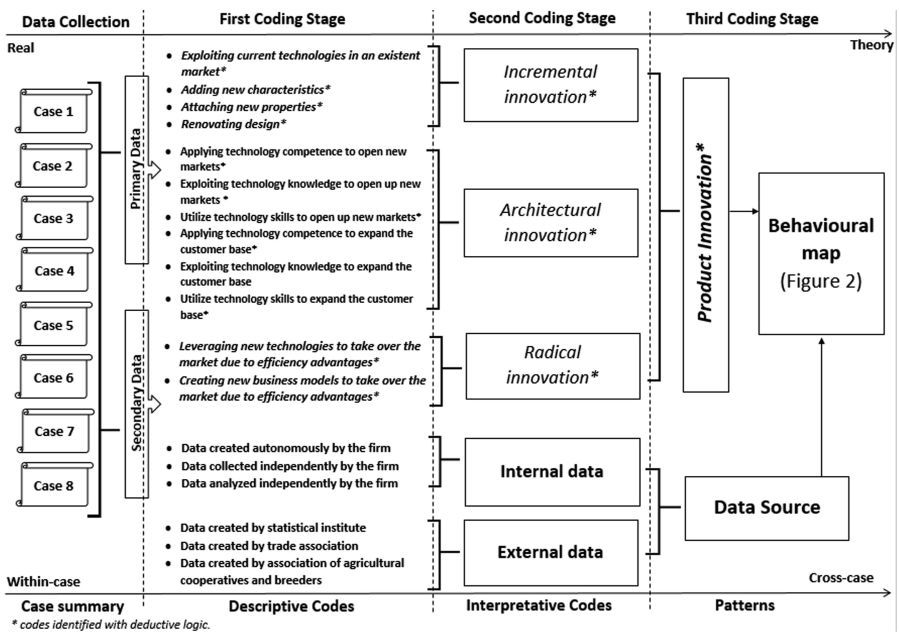
about product innovation? Which kind of analysis was more helpful for your company in the process of product innovation?

We recorded and wrote out the eight interviews in full within twenty-four hours. We collected the interviews from February 2023 to April 2023, and the duration of the interviews was from forty-two to fifty-eight minutes. We integrated the interviews by gathering data from companies' social media, websites, and, when provided, internal plans and reports to triangulate data sources. We evaluated the data collection using theoretical saturation methodology (Saunders *et al.*, 2018). We met data saturation in the eighth case.

3.3 Data analysis

We led data analyses in four rounds of coding with the support of NVivo 10 software (Cabiddu *et al.*, 2018). We employed both inductive and deductive logic in data analysis. Thus, we exploit earlier literature in a deductive way to interpret how agri-food firms employ digital technologies for product innovation (see the codes with the * in Figure 1 and Table 2). Also, we identified and formalised new theoretical constructs and related links inductively (Kennedy and Thornberg, 2018). Then, we began with an initial within-case analysis of the eight cases and their features by creating case summaries (Saldaña, 2015).

In the first coding round, we segmented and clustered data following a data-driven coding scheme during the first coding round. As an outcome, we pinpointed a list of descriptive codes as observed in the single-considered cases (Miles and Huberman, 1994). In the second coding round, we started the abstraction process by classifying new data under current codes, joining analogous codes or generating a new code if it was analytically different. Therefore, we analysed again the descriptive codes seeking interpretative codes (Miles and Huberman, 1994) (see Table 2).



Source(s): Own elaboration

Figure 1.
Data analysis process

Descriptive code	Definition	Illustrative quote	Interpretative code	Definition	Illustrative quote
Incremental innovation*	Innovation that exploits existing technologies in an existent market to revamp the current offering by adding new characteristics, properties and design	"The inclusion of new ingredients allows you to create a product that is not only tastier but also with a more constant quality and longer shelf life of the product. Certainly, this product innovation activity has been of great help to us. These innovations are not a product revolution, but they improve existing products." *Case study 5*	Product innovation*	The design of products with novel or incremented features that provide new or more customer benefits or are launched to a new market	"We got several advantages from data analysis. For instance, we have owned shops where we sell our products. There we collect data related to customers' reactions when we want to test new product or variances of the original product." *Case study 5*
Architectural innovation*	Innovation that takes technology competence, knowledge, and skills to apply to open up new markets and expand customer base	"A few years ago, I started making ginger products because I realized that ginger had a marketing appeal. Everyone thought that ginger was the solution to all ills. I just added an ingredient that none of my competitors used. The products have had great sales success." *Case study 2*			
Radical innovation*	Innovation that leverages new technologies, processes, or business models to take over the market due to efficiency and/or efficacy advantages	"In the dairy we have implemented a new technology that allows us to make mozzarella almost at industrial levels while maintaining an artisanal process that allows us to give a different flavor and a different texture to our mozzarella. Thus, it allows us to maintain a taste that industrial production does not guarantee." *Case study 4*			

(*continued*)

Table 2.
Summary of the descriptive and interpretative codes

Descriptive code	Definition	Illustrative quote	Interpretative code	Definition	Illustrative quote
Internal data	Data was created autonomously and independently by the agri-food firm	"I use Google Analytics on our website. Based on fresh pasta delivery requests, we collect addresses and it gives us an idea of where we might open other shop." *Case study 8*	Data sources	The data source discerns the data origin that, in turn, is used for different tasks and processes	"We analysis the sell trend of our performing products to improve their characteristic. Also, we gather data from external sources to do more extensive innovations." *Case study 6*
External data	Data created by a third party (e.g., statistical institute, trade association) and which are of interest to an agri-food firm	"to find an indirect customer need, the best way is to try to interpret the sectoral market data. Market data are provided by the trade association and are national data." *Case study 7*			

Table 2.

Source(s): Authors' own elaboration

The third coding round allowed us to identify patterns starting from the previously identified interpretative codes (Miles and Huberman, 1994). We matched data from multiple cases and verified whether a pattern was distinguishing only a case or recurrent in numerous cases. Consequently, the multiple cases worked as replication logic for the findings of this research, as contrary replication (detecting cases where the disposal of digital data did not result in product innovation) or as the exclusion of alternative justifications (discovering a different explanation for product innovation) (Eisenhardt and Graebner, 2007) (see Table 3). As a result, we reached a further level of abstraction. We pinpointed six behaviours that make the behavioural map (Figure 2) explaining how digital data are employed for product innovation: the supervisor, the passive supervisor, the developer, the passive developer, the conjunction behaviour, and the pathfinder.

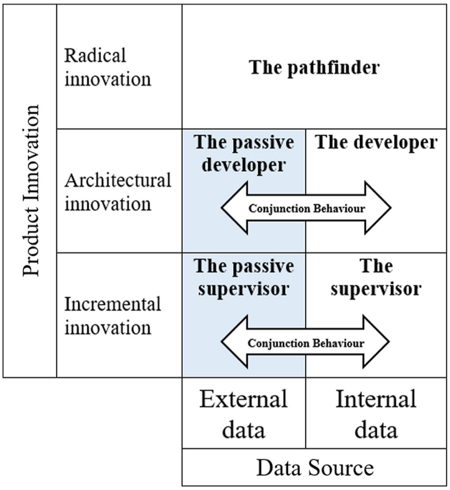
In the final coding round, data analysis required identifying the connections between patterns. This stage aims to link the patterns and convert them into a dynamic and cohesive theoretical framework from stationary and separate behaviours. We intersect the descriptive and interpretative codes (see Figure 2).

Behaviour \ Case Study	1	2	3	4	5	6	7	8
Supervisor			**X**			X	**X**	
Passive supervisor	X		**X**				**X**	X
Developer		**X**			**X**			X
Passive developer		**X**		X	**X**			
Pathfinder	X			X		X		

*Cells in bold represent the conjunction behaviour.

Source(s): Own elaboration

Table 3.
Cross-case summary of the behaviours



Source(s): Own elaboration

Figure 2. Behavioural map of data sources utilisation for product innovation

The co-authors sorted the emerging codes and their relationships to make the coding process more robust. During each round, the coders independently and separately analysed data and harmonised their code taxonomy by running a *coding comparison query*. The coders debated about the classification and solved inconsistencies until the value of the k coefficient was above 0.75, which Landis and Koch (1977) considered a substantial degree of agreement or even very good following Fleiss *et al.* (2003) classification.

4. Findings

4.1 How agri-food firms utilise digital data for product innovation

There are many reasons why agri-food companies analyse data. This article only studies digital data usage for product development in terms of different kinds of innovation. Agri-food companies create, collect and analyse data from internal sources such as raw materials, production processes, and warehouses. In addition, agri-food firms may also gather and analyse external data related to market trends, partnerships, competitors, and overall sector sales. We compared different innovation processes with the data sources to identify how agri-food companies behave when utilising digital data sources for product innovation. We found six behaviours connected to digital data utilisation described in the following paragraphs. We also observed that some of these behaviours could be seen simultaneously in the same agri-food company (conjunction behaviour). The findings about digital data utilisation behaviours and their relationships are graphically displayed in Figure 2.

4.2 The roadmap to the agri-food companies' data utilisation behaviours

4.2.1 The passive supervisor. The passive supervisor waits for prearranged sectorial data to check whether the market is ready for an incremental-innovated product. Therefore, the passive supervisor has an inert attitude to data analysis, mainly done on external data to produce information to foster incremental innovations. Consistent with our findings, the passive supervisors analyse data concerning internal aspects of the firm to improve current products by exploiting internally created data. Passive supervisors aim to progress their current products or create variations, changing some of their characteristics. To do that, the passive supervisors analyse their external environments, such as sales or mark trends: "*We track everything we sell. We usually invest in products that sell the most to create variations. The market will likely accept these variations"* Case study 3. Still, passive supervisors can employ external data to create new food products. For example, agri-food firms which are passive supervisors may analyse competitors' products to emulate them: "*We are followers regarding creating new products. We observe large companies that can make essential investments in research and development . . . then we analyse their products and try to adapt to what the most prominent companies do"* Case study 1.

4.2.2 The supervisor. The supervisor creates and collects internal production data to estimate incremental innovation feasibility and costs. Thus, the supervisors examine internal data to feed the innovation of their products. According to our findings, the supervisors analyse data concerning internal aspects of the firm, such as internally-gathered production data to innovate their current products incrementally: "*Each production is a test. The data are analysed to identify the strengths and weaknesses of the product. Over time, we improve our products, for example, by changing the flour mixture or inserting another type of raw material"* Case Study 7. Data are used to estimate the production costs and to understand whether the production plant is capable of producing a similar product but which has new characteristics that make the products more interesting for the final customer: *"We have accounting software that, based on the business cost structure and the recipe we want to give to the new product, can estimate the cost per pack. The software has all the parameters, such as production cost, labour cost, energy cost, raw material cost, cartons cost, label cost, etc. When we estimate the cost per pack, we also know if the new product is competitive."* Case study 6. Therefore, the supervisors use internal data mainly to estimate the production plant's innovation feasibility and cost-effectiveness.

4.2.3 The passive developer. The passive developer collects external information about new markets to assess the possibility of satisfying new needs with the current products. In the more straightforward scenario, the customer makes a specific request: *"The large-scale distribution sometimes says they might need a new product. For example, they say that since the market is always going towards vegan, is it possible to have tortellini with vegan dough? Then from this input, we try to satisfy this need with what we already produce."* Case study 7. Digital platforms such as review portals help agri-food companies collect information about new market opportunities: *"From the point of view of the review pages on the Internet, I would say that what we have exploited the most over time has been the data that we received on the types of food such as pasta shapes and sauces. Very often we rely on what we read in the reviews. Since most of the information, we get from there, from the reviews."* Case study 8.

4.2.4 The developer. The developer analyses the firm's technological competencies, knowledge, and skills to apply them to new markets. However, other stakeholders often influence this behaviour, especially the technology suppliers, which may limit the agri-food capability to enter new markets: *"The technology supplier is always quite reluctant to modify the technology to suit our needs of breaking into a new market. It takes a lot of effort. In our case, the technology supplier is a monopolist and imposes their solutions, and this is a limit. Then the supplier of the systems proposes the relative software components which collect the data and carries out its processing which is entirely centred on the functioning of the machine. However, the problem is that much other information could be used for other reasons"* Case study 6.

On the other hand, the technology suppliers can also be allies, particularly for small and medium-sized agri-food companies: *"The most important technological transformation has been the possibility of purchasing machinery that produces small volumes while maintaining the same performance. These plants create costs proportional to the production capacity. Our company use the latest technologies to adapt to the market dynamics to meet the demand in terms of quality and keep up with the market needs."*

4.2.5 The conjunction behaviour. When innovating their products, agri-food companies committed to incremental and architectural innovation can conjunctively behave as supervisors, passive supervisors or developers and passive developers in distinct situations. Firms in these groups behave as passive when they exploit external data to make simple or routinised decisions (see Figure 2). In so doing, conjunction behaviour companies exploit the standard information generated by, for example, their information system. Conversely, when conjunction behaviour companies face new problems in their innovation process and need to make uncommon decisions, they also need to search for external data to find information to help the decision-making process.

4.2.6 The pathfinder behaviour. The pathfinder proactively examines external and internal digital data to generate information to support the radical innovation process. This means that the pathfinder behaviour is taken on by the agri-food firms that explore digital data by analysing them with a critical eye to find novel pieces of information. Digital data can be both external and internal data sources. Our analysis pinpointed a few pathfinders that inquiry their database to create radically new products and understand how they can make them. As detected in Case Study 1*, production data are analysed to create new products. Therefore, we open new markets by analysing the production data to understand if the production plant can produce a product that presents new characteristics that make the product more attractive to the final customer. For example, it happened when we created the hamburger made of seitan.*

Nevertheless, the pathfinders often employ a combination of the data sources, as detected in Case 4: *"Cross-referencing data is useful. The general data of the market are at a national level and monthly. They are provided to us by the trade association. By cross-referencing our data with general trends, we can understand that a product that does not have large sales volumes nationwide is a product that works for our company. This could indicate that the product indirectly satisfies a consumer need."* Our analysis finds that pathfinder is the behaviour that better exploits the disposal of digital data since this behaviour pushes agri-food companies to employ internal and external data to perform radical innovation.

5. Discussion

Using qualitative empirical data from a multiple-case study, this study seeks to answer the following research question: How do agri-food companies employ various data sources to drive product innovation?

The findings uncover six different patterns of digital data use for different types of agri-food product innovations: supervisor, passive supervisor, developer, passive developer, conjunction behaviour, and pathfinder. In our research, we distinguish three types of product innovations: incremental, architectural and radical. External and internal digital data may play a role in each of these types of innovation, but the patterns of use and how the data are combined differ. For incremental innovation, one approach involves the supervisor generating and gathering internal production data to assess the feasibility and costs of such innovation. In contrast, the passive supervisor relies on predetermined sectorial data to gauge market readiness for an incrementally innovative product. Consequently, the passive supervisor demonstrates a less engaged stance towards data analysis, primarily relying on external data to facilitate incremental innovations. In the realm of architectural innovation, developers evaluate the firm's technological capabilities, knowledge, and expertise to apply them to new markets. Conversely, passive developers collect external data on new markets to evaluate the potential for meeting new demands with existing products. Also, when dealing with incremental and architectural innovation, agri-food companies exhibit conjunctive behaviours, acting as supervisors, passive supervisors, developers, or passive developers in different scenarios. Firms tend to be passive when utilising external data to make routine or straightforward decisions. Conversely, they actively generate and analyse data when encountering new challenges in the product innovation process, necessitating unconventional decisions. Lastly, pathfinders proactively analyse and integrate internal and external digital data to generate insights supporting radical innovation. This behaviour stands apart from conjunction because agri-food firms cannot effectively undertake radical innovation without integrating both internal and external data sources.

5.1 Theoretical implications

Our study contributes to the recent academic debate on the role of digital technologies in agri-food product innovation. In the wake of the recent proliferation of digital technologies, studies have examined the impact of different digital technologies on innovation. The diversity of digital technologies under scrutiny, such as precision technologies, machine learning, drones and sensors for digital data generation, are just a few examples (Oltra-Mestre *et al.*, 2021; Romanello and Veglio, 2022). Rather than examining the impact of a specific digital technology on a specific innovation, our study introduces a new analytical perspective to the literature on the topic, distinguishing between external and internal digital data. Our study extends previous research by providing classification and explanation of digital data behaviours in the agri-food industry and by showing how agri-food companies can effectively use different digital data sources to support a variety of food product innovations.

This study sheds light on the underlying patterns of how firms use and integrate digital data from different sources. The notion of incorporating external data and information into the process of successful innovation is not novel in the innovation literature. Whilst most studies emphasise the pivotal role of using external data (e.g., Du, 2021; Li and Tamer Cavusgil, 2000; Zhang and Xiao, 2020), less attention has been paid to the role of internal data and the importance of combining external and internal data sources. This study concludes that external data does not always pay off because it can lead to a passive attitude towards data utilisation, like the passive developer and supervisor for architectural and incremental innovation. Firms need a conjunction behaviour to be more active in creating and analysing data related to inside phenomena, for example, by installing sensors in the production plan to feed unexpected architectural and incremental innovation challenges. The most prominent example is the pathfinder, which cannot avoid integrating both internal and external data sources to face radical innovation successfully.

Finally, our study contributes to a more nuanced and comprehensive understanding of how firms combine digital data in case of different types of product innovation. Although digital technology scholars from the field of agri-food empirically validate the link between digital data utilisation and innovation outcomes (Zambon *et al.*, 2019; Demartini *et al.*, 2018), they pay less attention to the underlying mechanism and do not show how different data sources lead to different innovation outcomes leaving the academic debate with a flat view of data utilisation for product innovation. Outside the realm of agri-food literature, we found studies exploring the connections between data sources and product innovation outcomes. These studies present conflicting findings. For instance, Ritala *et al.* (2018) suggested that indiscriminate external data usage reduced performance in radical innovation with less impact on incremental innovation. Other studies advised the opposite as they found a positive relation between external data sources and radical innovation (e.g., Maes and Sels, 2014; Du, 2021). Finally, some scholars claim that there is a positive impact of internal data sources and incremental product innovation (e.g., Du, 2021; Hutchinson, 2020). Nevertheless, prior literature not only did not focus on the agri-food industry but also did not systematically investigate how companies utilise internal and external data sources for incremental, architectural and radical innovation, providing a fragmented picture of the phenomenon.

5.2 Managerial implications

To maximise the potential of digital data for product innovation, agri-food companies should adopt a proactive approach to analysing internal and external data. For example, they could utilise prearranged external data analysis services to gain insights into consumer trends and preferences, which could inform decisions about incremental innovation on current products. Similarly, they could perform routinised analyses of external data, such as market research and competitor data, to identify opportunities for architectural innovation, like using new technologies to enter new markets and expand their customer base.

However, when it comes to analysing internal data, agri-food companies must change their approach to identifying correlations amongst food processing data and identifying opportunities for radical innovation. For instance, they could analyse data from sensors installed in their production lines to identify patterns and anomalies that could indicate new product opportunities or production efficiencies. In addition, pathfinder firms that effectively leverage internal and external data can perform radical innovation. For example, a food manufacturer could use customer reviews and feedback data to identify unmet needs in the market and create new products to meet those needs.

Agri-food companies that combine supervisory and developer behaviours can better exploit their data sources to feed their incremental or architectural innovation efforts. For instance, they could establish cross-functional teams that include data scientists, product developers, and business leaders to identify and prioritise data-driven opportunities for innovation. Therefore, agri-food companies interested in maximising the potential of their digital data should prioritise a proactive and flexible approach to data analysis and leverage internal and external data to drive innovation.

5.3 Limitations and future research

This study's qualitative exploratory multiple case study methodology involves limits that offer empirical and theoretical investigation opportunities. First, we collected data from only food processing firms. Thus, our study's empirical setting gives a fractional perspective of the transformation due to digital technologies in the wider agri-food industry. Moreover, we gathered data from a sample of eight cases located in the Italian and Hungarian markets. Therefore, we suggest further research to spread our methodology to other agri-food industries and cultural environments. In particular, we call for more research in agriculture and retail because of the fast and greater use of digital tools.

Moreover, we suggest extending this study to international and culturally diverse contexts. The international and multicultural environment will allow us to verify whether the undefined behaviours change depending on the context. Hence, such a manifold research setting may help pinpoint new behaviours driven by using digital data for product innovation.

Additionally, this study takes the manufacturing perspective. Therefore, we based the data analysis and showed the findings considering only one actor. Nevertheless, using digital data for product innovation involves many stakeholders in the food sector. Furthermore, each stakeholder may affect the utilisation of digital data for product innovation differently. Thus, future research might engage several actors when collecting data to provide multiple perspectives when investigating the same phenomena. For example, future research might involve dedicated agri-food machinery suppliers and software houses which develop digital tools (e.g., applications and equipment) for farms, food manufacturers, and retailers.

Ultimately, the proposed theoretical framework, the behavioural map, could gain value through quantitative validation and testing. Subsequent research endeavours could involve the formal creation of measurement scales for the six identified behaviours alongside the validation of a suitable survey instrument for their assessment (MacKenzie *et al.*, 2011).

Such an approach would furnish researchers with a robust scientific instrument, facilitating explanatory research and enabling exploration of this topic across various organisational contexts (Straub, 1989).

References

Adamashvili, N., Fiore, M., Contò, F. and La Sala, P. (2020), "Ecosystem for successful agriculture. Collaborative approach as a driver for agricultural development", *European Countryside*, Vol. 12 No. 2, pp. 242-256, doi: 10.2478/euco-2020-0014.

Adamashvili, N., State, R., Tricase, C. and Fiore, M. (2021), "Blockchain-based wine supply chain for the industry advancement", *Sustainability*, Vol. 13 No. 23, 13070, doi: 10.3390/su132313070.

Beckeman, M., Bourlakis, M. and Olsson, A. (2013), "The role of manufacturers in food innovations in Sweden", *British Food Journal*, Vol. 115 No. 7, pp. 953-974, doi: 10.1108/BFJ-09-2010-0164.

Belaud, J.P., Prioux, N., Vialle, C. and Sablayrolles, C. (2019), "Big data for agri-food 4.0: application to sustainability management for by-products supply chain", *Computers in Industry*, Vol. 111, pp. 41-50, doi: 10.1016/j.compind.2019.06.006.

Ben Ayed, R., Hanana, M., Ercisli, S., Karunakaran, R., Rebai, A. and Moreau, F. (2022), "Integration of innovative technologies in the agri-food sector: the fundamentals and practical case of DNA-based traceability of olives from fruit to oil", *Plants*, Vol. 11 No. 9, p. 1230, doi: 10.3390/plants11091230.

Büchi, G., Cugno, M. and Castagnoli, R. (2020), "Smart factory performance and industry 4.0", *Technological Forecasting and Social Change*, Vol. 150, 119790, doi: 10.1016/j.techfore.2019.119790.

Cabiddu, F., Frau, M. and Moi, L. (2018), "Exploring the role of NVivo software in marketing research", *Mercati e Competitività*, No. 4, pp. 65-86, doi: 10.3280/MC2018-004005.

Danneels, E. (2002), "The dynamics of product innovation and firm competences", *Strategic Management Journal*, Vol. 23 No. 12, pp. 1095-1121, doi: 10.1002/smj.275.

Demartini, M., Pinna, C., Tonelli, F., Terzi, S., Sansone, C. and Testa, C. (2018), "Food industry digitalisation: from challenges and trends to opportunities and solutions", *IFAC-PapersOnLine*, Vol. 51 No. 11, pp. 1371-1378, doi: 10.1016/j.ifacol.2018.08.337.

Du, L. (2021), "How knowledge affects incremental innovation in SMEs: knowledge base and knowledge heterogeneity", *Journal of General Management*, Vol. 46 No. 2, pp. 91-102, doi: 10.1177/0306307020930196.

Eisenhardt, K.M. and Graebner, M.E. (2007), "Theory building from cases: opportunities and challenges", *Academy of Management Journal*, Vol. 50 No. 1, pp. 25-32, doi: 10.5465/amj.2007.24160888.

Endres, C.M., Pelisser, C., Finco, D.A., Silveira, M.S. and Piana, V.J. (2022), "IoT and Raspberry Pi application in the food industry: a systematic review", *Research, Society and Development*, Vol. 11 No. 1, e0411124270, doi: 10.33448/RSD-V11I1.24270e0411124270-e0411124270.

European Commission (2023), "European innovation scoreboard", available at: https://research-and-innovation.ec.europa.eu/statistics/performance-indicators/european-innovation-scoreboard_en (accessed 8 February 2024).

Ferraris, A., Mazzoleni, A., Devalle, A. and Couturier, J. (2019), "Big data analytics capabilities and knowledge management: impact on firm performance", *Management Decision*, Vol. 57 No. 8, pp. 1923-1936, doi: 10.1108/MD-07-2018-0825.

Fleiss, J., Levin, B. and Paik, M. (2003), *Statistical Methods for Rates & Proportions*, 3rd ed., Wiley & Sons, New York.

Frau, M., Moi, L. and Cabiddu, F. (2020), "Outside-in, inside-out, and blended marketing strategy approach: a longitudinal case study", *International Journal of Marketing Studies*, Vol. 12 No. 3, pp. 1-13, doi: 10.5539/ijms.v12n3p1.

Frau, M., Moi, L. and Cabiddu, F. (2022a), "Digital transformation through the lens of digital data handling: an exploratory analysis of agri-food SMEs", *Journal of Small Business Strategy*, Vol. 32 No. 3, pp. 84-97, doi: 10.53703/001c.34642.

Frau, M., Moi, L., Cabiddu, F. and Keszey, T. (2022b), "Time to clean up food production? Digital technologies, nature-driven agility, and the role of managers and customers", *Journal of Cleaner Production*, Vol. 377, 134376, doi: 10.1016/j.jclepro.2022.134376.

Hanelt, A., Bohnsack, R., Marz, D. and Antunes Marante, C. (2021), "A systematic review of the literature on digital transformation: insights and implications for strategy and organisational change", *Journal of Management Studies*, Vol. 58 No. 5, pp. 1159-1197, doi: 10.1111/joms.12639.

Hassoun, A., Boukid, F., Pasqualone, A., Bryant, C.J., García, G.G., Parra-López, C., Barba, F.J., Trollman, H. and Cropotova, J. (2022), "Emerging trends in the agri-food sector: digitalisation and shift to plant-based diets", *Current Research in Food Science*, Vol. 5, pp. 2261-2269, doi: 10.1016/j.crfs.2022.11.010.

Hutchinson, P. (2020), "Reinventing innovation management: the impact of self-innovating artificial intelligence", *IEEE Transactions on Engineering Management*, Vol. 68 No. 2, pp. 628-639, doi: 10.1109/TEM.2020.2977222.

Kamble, S.S., Gunasekaran, A. and Gawankar, S.A. (2020), "Achieving sustainable performance in a data-driven agriculture supply chain: a review for research and applications", *International Journal of Production Economics*, Vol. 219, pp. 179-194, doi: 10.1016/j.ijpe.2019.05.022.

Kennedy, B.L. and Thornberg, R. (2018), *Deduction, Induction, and Abduction*, The SAGE Handbook of Qualitative Data Collection, London, pp. 49-64.

Keszey, T. (2018), "Boundary spanners' knowledge sharing for innovation success in turbulent times", *Journal of Knowledge Management*, Vol. 22 No. 5, pp. 1061-1081, doi: 10.1108/JKM-01-2017-0033.

Konfo, T.R.C., Djouhou, F.M.C., Hounhouigan, M.H., Dahouenon-Ahoussi, E., Avlessi, F. and Sohounhloue, C.K.D. (2023), "Recent advances in the use of digital technologies in agri-food processing: a short review", *Applied Food Research*, Vol. 3 No. 3, 100329, doi: 10.1016/j.afres.2023.100329.

Kör, B., Krawczyk, A. and Wakkee, I. (2022), "Addressing food loss and waste prevention", *British Food Journal*, Vol. 124 No. 8, pp. 2434-2460, doi: 10.1108/BFJ-05-2021-0571.

Laforet, S. (2009), "Effects of size, market and strategic orientation on innovation in non-high-tech manufacturing SMEs", *European Journal of Marketing*, Vol. 43 Nos 1/2, pp. 188-212, doi: 10.1108/03090560910923292.

Landis, J.R. and Koch, G.G. (1977), "The measurement of observer agreement for categorical data", *Biometrics*, Vol. 33 No. 1, pp. 159-174, doi: 10.2307/2529310.

Lezoche, M., Hernandez, J.E., Díaz, M.D.M.E.A., Panetto, H. and Kacprzyk, J. (2020), "Agri-food 4.0: a survey of the supply chains and technologies for the future agriculture", *Computers in Industry*, Vol. 117, 103187, doi: 10.1016/j.compind.2020.103187.

Li, T. and Tamer Cavusgil, S. (2000), "Decomposing the effects of market knowledge competence in new product export: a dimensionality analysis", *European Journal of Marketing*, Vol. 34 Nos 1/2, pp. 57-80, doi: 10.1108/03090560010306205.

MacKenzie, S.B., Podsakoff, P.M. and Podsakoff, N.P. (2011), "Construct measurement and validation procedures in MIS and behavioral research: integrating new and existing techniques", *MIS Quarterly*, Vol. 35 No. 2, pp. 293-334, doi: 10.2307/23044045.

Maes, J. and Sels, L. (2014), "SMEs' radical product innovation: the role of internally and externally oriented knowledge capabilities", *Journal of Small Business Management*, Vol. 52 No. 1, pp. 141-163, doi: 10.1111/jsbm.12037.

Matzler, K., Friedrich von den Eichen, S., Anschober, M. and Kohler, T. (2018), "The crusade of digital disruption", *Journal of Business Strategy*, Vol. 39 No. 6, pp. 13-20, doi: 10.1108/JBS-12-2017-0187.

Miles, M.B. and Huberman, A.M. (1994), *Qualitative Data Analysis: an Expanded Sourcebook*, Sage, Thousand Oaks.

Müller, J.M., Buliga, O. and Voigt, K.I. (2020), "The role of absorptive capacity and innovation strategy in the design of industry 4.0 business models—a comparison between SMEs and large enterprises", *European Management Journal*, Vol. 39 No. 3, pp. 333-343, doi: 10.1016/j.emj.2020.01.002.

Musa, SFPD and Basir, K.H. (2021), "Smart farming: towards a sustainable agri-food system", *British Food Journal*, Vol. 123 No. 9, pp. 3085-3099, doi: 10.1108/BFJ-03-2021-0325.

Oltra-Mestre, M.J., Hargaden, V., Coughlan, P. and Segura-García del Río, B. (2021), "Innovation in the agri-food sector: exploiting opportunities for industry 4.0", *Creativity and Innovation Management*, Vol. 30 No. 1, pp. 198-210, doi: 10.1111/caim.12418.

Piccoli, G. and Watson, R.T. (2008), "Profit from customer data by identifying strategic opportunities and adopting the 'born digital' approach", *MIS Quarterly Executive*, Vol. 7 No. 3, pp. 113-122.

Ritala, P., Husted, K., Olander, H. and Michailova, S. (2018), "External knowledge sharing and radical innovation: the downsides of uncontrolled openness", *Journal of Knowledge Management*, Vol. 22 No. 5, pp. 1104-1123, doi: 10.1108/JKM-05-2017-0172.

Romanello, R. and Veglio, V. (2022), "Industry 4.0 in food processing: drivers, challenges and outcomes", *British Food Journal*, Vol. 124 No. 13, pp. 375-390, doi: 10.1108/BFJ-09-2021-1056.

Saldaña, J. (2015), *The Coding Manual for Qualitative Researchers*, Sage, London.

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H. and Jinks, C. (2018), "Saturation in qualitative research: exploring its conceptualisation and operationalisation", *Quality and Quantity*, Vol. 52 No. 4, pp. 1893-1907, doi: 10.1007/s11135-017-0574-8.

Schweitzer, F.M., Handrich, I.M. and Heidenreich, S. (2019), "Digital transformation in the new product development process: the role of it-enabled PLM systems for relational, structural, and NPD performance", *International Journal of Innovation Management*, Vol. 23 No. 7, 1950067, doi: 10.1142/S1363919619500671.

Steenis, N.D. and Fischer, A.R. (2016), "Consumer attitudes towards nanotechnology in food products: an attribute-based analysis", *British Food Journal*, Vol. 118 No. 5, pp. 1254-1267, doi: 10.1108/BFJ-09-2015-0330.

Straub, D.W. (1989), "Validating instruments in MIS research", *MIS Quarterly*, Vol. 13 No. 2, pp. 147-169, doi: 10.2307/248922.

Varese, E. and Cane, P. (2017), "From space food research and innovation to immediate advantages for Earth eating habits: an aerospace–food producer company case study", *British Food Journal*, Vol. 119 No. 11, pp. 2448-2461, doi: 10.1108/BFJ-04-2017-0216.

Yin, R.K. (2009), *Case Study Research: Design and Methods*, 4th ed., Library of Congress Cataloguing-In-Publication Data, United States.

Zambon, I., Cecchini, M., Egidi, G., Saporito, M.G. and Colantoni, A. (2019), "Revolution 4.0: industry vs. Agriculture in a future development for SMEs", *Processes*, Vol. 7 No. 1, p. 36, doi: 10.3390/pr701003.

Zhang, H. and Xiao, Y. (2020), "Customer involvement in big data analytics and its impact on B2B innovation", *Industrial Marketing Management*, Vol. 86, pp. 99-108, doi: 10.1016/j.indmarman.2019.02.020.

Corresponding author
Moreno Frau can be contacted at: moreno.frau@uni-corvinus.hu