Technological Forecasting & Social Change 192 (2023) 122515

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

Data sharing for business model innovation in platform ecosystems: From private data to public good

Nikolai Kazantsev [a], Nazrul Islam [b,*], Jeremy Zwiegelaar [c], Alan Brown [d], Roger Maull [d]

[a] Institute for Manufacturing, University of Cambridge, UK
[b] Royal Docks School of Business and Law, University of East London, UK
[c] Oxford Brookes Business School, Oxford Brookes University, UK
[d] University of Exeter Business School, UK

ARTICLE INFO

Keywords:
Open data
Business model innovation
Platform ecosystems
Public good

ABSTRACT

Extant research posits that open data could unlock more than $3 trillion in additional value worldwide across various application domains. This paper investigates a data-sharing perspective in business models of platform ecosystems and discusses how platform owners can derive more value using data. We chose a sample of 12 platforms in which data are used as a key resource for service propositions. By contrasting these cases, we identify and analyse four archetypes: data crawler, data marketplace, data aggregator, and data disseminator. We define the key features of these archetypes and demonstrate how they realise value via the platform. These archetypes can guide managers in realising private and public goods via data sharing. Building on our findings, we derive recommendations for data-driven business model innovation for platform ecosystems.

1. Introduction

Companies such as Amazon, Microsoft, IBM, Google, and Facebook have mastered vast, interconnected computing utilities across the globe to explore data generated by billions of consumers and gain leadership in their markets (Brown, 2021). Their consumers gave them access to petabytes of private data, attracted by ubiquitous access to services through internet-connected mobile devices such as tablets, smartphones, internet of things (IoT) sensors, and an array of wearable technologies. Extant studies predict that open data will unlock more than $3 trillion in additional value worldwide across various application domains (Candelon et al., 2020; Manyika et al., 2013); thus, clarifying how data sharing can support widespread impacts (European Commission, 2016; Hampton et al., 2015) using a business model innovation lens (Teece, 2018) presents an intriguing research topic.

Both private and public data may be used for data-sharing purposes. However, the literature is unclear on how to increase the number of actors to capture the data-driven value (Magalhaes and Roseira, 2020; Zuiderwijk et al., 2014), considering that most actors sharing data do not get value in return. In particular, the literature is silent about how to increase data sharing for shared societal benefits. Society would see value from answering some of the biggest questions of our age; for example, will the connected car enable us to reduce congestion in cities and avoid accidents? Does greater insight into energy consumption via smart metering change consumers' behaviour to drive a more sustainable approach to energy management? Does the adoption of wearable health monitors lead to earlier interventions to increase wellness and ensure a longer, more active life in consumers' advancing years at an affordable price? (Brown, 2021). Answering these questions requires more research on the impact of data sharing beyond their economic effect on the firm (Snihur and Bocken, 2022), also referred to as 'public goods' by economists (Stiglitz, 1999). Using a business model innovation lens (i.e., value creation, delivery, and capture), this paper explores how to realise more via data embedded in a business or social context using the data set of platform ecosystems enabled by data sharing (Foss and Saebi, 2017; Teece, 2018; Yoo et al., 2010) (see Fig. 1).

Our research question (RQ) is as follows: *What are the business model archetypes of platform ecosystems from a data-sharing perspective?* Answering this RQ will help large organisations looking for new ways to explore and exploit their data and the data of their suppliers and customers to which they have access. Our research can also help smaller firms that view data sharing as an important resource to gain more value. Finally, actors generating or holding data may find important messages regarding the application of their data assets. To answer our RQ, we completed a multiple-case study of 12 platform ecosystems enabled by data sharing to analyse how value is realised. First, we derive

* Corresponding author.
E-mail address: Nazrul.Islam@uel.ac.uk (N. Islam).

https://doi.org/10.1016/j.techfore.2023.122515
Received 15 September 2021; Received in revised form 7 December 2022; Accepted 2 January 2023
Available online 31 March 2023

four business model archetypes of platforms based on data sharing: data crawler, data marketplace, data aggregator, and data disseminator. In doing so, we extend the current knowledge of business model innovation via data sharing. Second, we explain how these archetypes function and evolve using a two-by-two classification, thus providing greater clarity on how value is derived (Janssen and Zuiderwijk, 2014; Magalhaes and Roseira, 2020). Third, we suggest a new mechanism for value realisation using these archetypes, thus contributing to the research on business models in general and value propositions in particular (Zuiderwijk et al., 2014).

2. Literature review

2.1. Data platform ecosystems

A digital platform provides standards, interfaces, and rules to support innovation embedded in the business or social context of a broader ecosystem (Yoo et al., 2010). Platform ecosystems comprise technological building blocks that the owner and partners can share, such as the *platform core*, which provides basic functionality to modular services (Tiwana et al., 2010); *boundary resources*, which provide interface and support to ecosystem actors (Teece and Linden, 2017); and *complements*, which are individual ecosystem components (Hein et al., 2020). Platform ecosystems are networks in which platform owners enable third parties to create complementary innovations (Cozzolino et al., 2021; Parida et al., 2019) and derive new products and services (Cusumano et al., 2019). For example, complementors can develop and co-create new apps to extend the platform's core functionality (Baldwin and Woodard, 2009) and commercialise these developments in the marketplace (Fruhwirth et al., 2020; Stahl et al., 2014). Data innovation ecosystems are a dynamic arrangement of actors that work together to share data and create value (Immonen et al., 2014). Adner (2006) underscores the role of customer centricity in the value proposition of innovation ecosystems: the more actors join the platform, the more opportunities arise to improve business using their data (Carnelley et al., 2013; Fruhwirth et al., 2020).

2.2. The impact of data on private and public goods

Data is a strategic resource in today's digital businesses (Bock and Wiener, 2017; Hartmann et al., 2016; Mamonov and Triantoro, 2018) 'digital fuel of the 21st century' (Kundra, 2012) and as an enabler of new economic activity and innovations (Davies and Perini, 2016). Data impacts all stages of value realisation, bridging a range of private and public goods. Private goods (e.g., financial income, products, services) are both rivalrous and excludable, whereas public goods are neither; in other words, '[goods] which all enjoy in common in the sense that each individual's consumption of such a good leads to no subtractions from any other individual's consumption of that good' (Samuelson, 1954).

As Stiglitz (1999) notes, public goods are noticed when the value capture exceeds the boundaries of a firm. For instance, a company selling IoT-enabled bike lights[1] can enable a platform between road users (i.e., citizens and cyclists) and public administration (i.e., city government) by improving awareness about city congestion and road quality. Likewise, data collaboratives facilitate private data sharing to enable their partners to find new, innovative, and data-driven solutions to public problems – from addressing climate change to public health to job creation (Verhulst et al., 2015). Society as a whole shares such a public good. To complement the existing business models with similar initiatives, we need an a priori conceptualisation of how a platform ecosystem would function in achieving its goals in a balance between profitability and social impact (Massa et al., 2017).

2.3. Data-driven business model innovation

At a very general and intuitive level, business models comprise a common set of components: value creation, value delivery, and value capture (Amit and Zott, 2001; Osterwalder and Pigneur, 2010; Teece, 2010). Open innovation (OI) is an example of data-driven *value creation* (Chesbrough and Rosenbloom, 2002; Hartmann et al., 2016; Wixom and Ross, 2017), in the sense that it rests on the principle that firms leverage internal and external ideas and paths to market to innovate, defining new organisational architectures and systems (Bogers et al., 2018). Chesbrough and Bogers (2014, p. 17) define OI as 'a distributed innovation process based on purposefully managed knowledge flows across organisational boundaries'. It involves harnessing inflows and outflows of knowledge to drive innovation within the firm and enable wider industrial impact (Davies and Perini, 2016). Digital platforms form the *value delivery* component of a data platform business model (Bonina et al., 2021). Firms that use data along the product life cycle trigger further changes in *value capture* (Bock and Wiener, 2017; Foss and Saebi, 2017; Otto and Aier, 2013). For example, contracting on outcomes

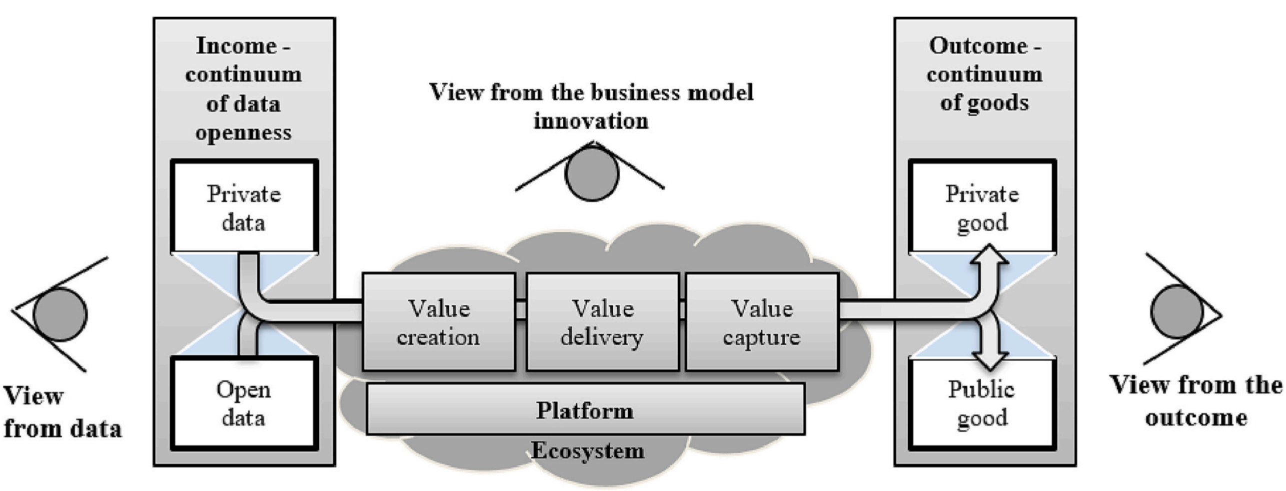

Fig. 1. Conceptual model of business model innovation in platform ecosystems as a convertor of data into private/public good.

that can redefine or even re-invent industry structures (Huikkola et al., 2020; Linde et al., 2021). The term 'open data' refers to data that are transferable (Manyika et al., 2013, p. 4); it has been described as the

[1] https://seesense.cc/ (accessed 9 August 2022).

requires access to data on how a customer uses the product. Outcome-based contracting is an advanced value capture mechanism for firms to focus on delivering value-in-use, as the firm would need to jointly deliver outcomes with the customer (Ng et al., 2010). For instance, Rolls-Royce uses private data to monitor jet engine utilisation rate and ensure availability and co-production results, such as the number of hours the engine is in the air. This type of data-driven contracting keeps the producer as product owner, who handles the risks associated with the product not in use (Demirkan and Spohrer, 2015).

2.4. Area of contribution

Much research on platform ecosystems has been inspired by data-sharing initiatives, such as Open Government Data programs (Adner, 2017). Bonina et al. (2021) point to open data innovation platforms as a promising research context in which to investigate business model innovation for development. However, we note the limited research on the purpose of business model innovation in platform ecosystems 'beyond the profits of an individual firm' (Snihur and Bocken, 2022). Although many companies are collecting data (Huyer and van Knippenberg, 2020), the literature is unclear on how to realise value for more actors (Magalhaes and Roseira, 2020; Zuiderwijk et al., 2014). We propose that the data owners are not incentivised to unlock their data (Foss and Saebi, 2017; Teece, 2018) when they neither understand the profit-oriented opportunities behind data sharing nor capture part of the data-driven value (Fruhwirth et al., 2020). Therefore, we see a need to systematise profit-oriented data-sharing opportunities (Mehta et al., 2021), which should lead to more data sharing on platform ecosystems and, ultimately, positive societal impact. Fig. 2 depicts the theoretical framework of this research.

3. Method

In line with Salzberger et al.'s (1999) recommended empirical research steps, we proceeded as follows: problem definition, data collection, data preparation, and data analysis. For our research, we chose a multiple case study approach (Eisenhardt, 1989) to identify commonalities and differences in how platform ecosystems use data for private or public good. This method is appropriate when collecting initial empirical evidence about relatively new phenomena (Miles and Huberman, 1994); it suits the exploratory nature of the research and accommodates small numbers of case studies (Eisenhardt and Graebner, 2007). To ensure the comparability of data, we adopted rigorous procedures for the case study selection, data collection, and analysis. We chose the data innovation platform as the unit of analysis, and the object of analysis is the data-driven value creation, delivery, and capture (Yin, 1993).

3.1. Problem definition and approach

The purpose of this research is theory building, not generalizability (Eisenhardt, 1989). Therefore, we first developed a unified research objective: to investigate business models of platform ecosystems from a data-sharing perspective, proposing that data sharing does not create value per se but rather provides the potential for value creation. Data sharing strengthens the value proposition of the data platform ecosystems, which enables the creation of new data-driven services; this in turn can help an organization target a wider group of customers, such as society or even the planet. Thus, by enabling data sharing, platform ecosystems complement the delivery of their private goods by delivering more public goods. An example of such public goods is a data-driven service offering to 'enrich' established products (Reim et al., 2015) and contribute to a sustainability agenda.

3.2. Case selection and sampling

In qualitative research, case selection must be carefully developed to generate meaningful results (Yin, 2009). For this reason, we considered domain-specific platforms (Thomas and Leiponen, 2016), which use data as the key resource in, for example, data science, public administration, antiviral software, and healthcare. We used a purposive sampling approach (Miles and Huberman, 1994) according to the following key criteria: (1) the platform had declared a goal of collecting data, (2) the platform aggregated and analysed the collected data sets, and (3) the platform provided data analysis results for business or research purposes. We collected qualitative secondary data from the data platforms during the period January–March 2021 using a desktop search, which resulted in identification of 31 data innovation platforms (for the final data set, see the Web Appendix, Table B[2]). Next, we evaluated each platform and agreed on commonalities and differences in value creation, delivery, and capture (Eisenhardt and Graebner, 2007). As described in Table 1, we limited the number of cases to the suggested 12 data innovation platforms, which are mostly different regarding data privacy levels and value creation, delivery, and capture mechanisms. Then, we identified the core platform characteristics – essential properties for building value from data (e.g., data sharing, data generation, data transformation).

3.3. Data analysis

To compare the cases, we created Microsoft Excel tables for each of the data platform categories; then, we summarised the findings in a final table in which we integrated and compared data from all the case studies. First, we searched for within-group similarities and intergroup differences to identify patterns in data privacy and value creation, delivery, and capture. Second, we carefully compared emergent frames against the evidence from each case to assess how well they fit the case data. This process led to the identification of cross-case patterns. Third, we validated findings against the literature by asking, 'what is this similar to, what does it contradict, and why'? (Yin, 1993). Using multiple cases increases the validity and reliability of findings, such as when research observations converge; ultimately, the confidence in findings improves (Eisenhardt, 1989; Eisenhardt and Graebner, 2007). Fourth, we ensured that we had reached the point of theoretical data saturation by noting when the data analysis provided no incrementally new insights into the research topic (Eisenhardt, 1989). Finally, the full author team checked the comparability among the cases and validated the final selection (see Fig. 3).

We adopted these rigorous procedures to increase the equivalence, reduce the bias, and establish the credibility and dependability of the results, which quantitative studies propose can be considered alternatives to reliability and validity (Guba and Lincoln, 1994; Sinkovics et al., 2008). Although we followed a procedure to select, collect, and analyse data, our study still represents an exploratory theory-building research project (Eisenhardt, 1989). Furthermore, the fact that the salient conditions overlap or match in different platform contexts demonstrates the transferability of the research findings, which represents an alternative way to measure generalisability in qualitative studies (Guba and Lincoln, 1994; Welch and Piekkari, 2017).

4. Results

The results obtained from our analysis can be clustered into the following four categories: *data crawler*, *data marketplace*, *data aggregator*, and *data disseminator* (see Table 1). They differ in terms of data privacy and business purpose – for instance, by capturing open data for a private good (crawlers), facilitating private data selling in market transactions

[2] This file is attached to the submission as a web appendix.

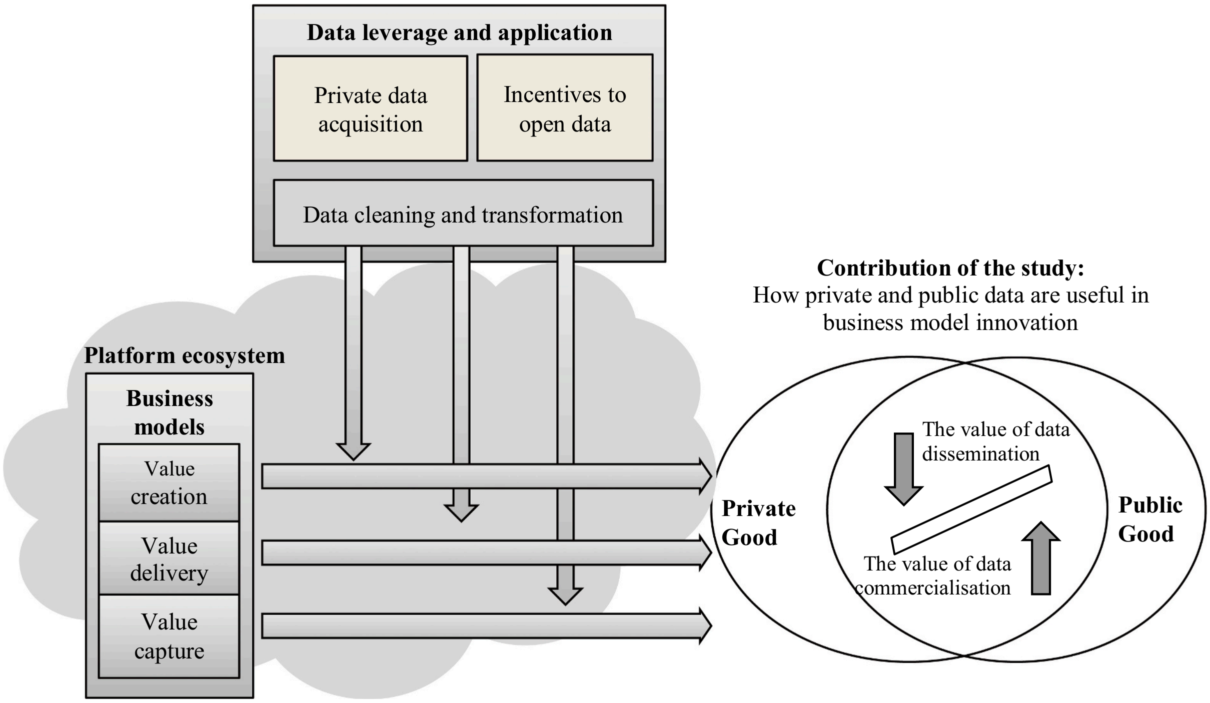

Fig. 2. The theoretical background and contribution of this study.

Table 1
Main characteristics of case study platforms.

Firm pseudonym	Platform name and website	Country	Industry	Data privacy	Data sharing incentive	Data generation	Data cleaning and transformation	Nature of good
Mercury	DataSift/ Meltwater	UK/USA	News and social media	Public	n/a	Web crawling	Enrichments, filtering, and classification	Private
Venus	Haix.AI	UK	Social media	Public	n/a	Web crawling	Cleaning and integration	Private
Earth	Pet Parade	USA	Social media/ pets	Public	n/a	Web crawling	Filtering and classification	Private
Mars	Breedr	UK	Animal health	Private	Data-based analytics	App	Predictive analytics	Private
Jupiter	Health Wizz	USA	Medical/ pharma	Private	Ease of data access	App	Data integration, classification, and analysis	Public
Saturn	Zenome	Russia	Biomedical	Private	Payment	App	Privacy management	Private
Neptune	Insilico Medicine	China	Biomedical	Private	Acceleration of drug discovery	Crowdsourcing	Cleaning and integration	Public
Uranus	See.Sense	UK	Smart city	Private	Improvement in traffic flow	Crowdsourcing	Data aggregation	Public
Pluto	Virustotal	Spain/ USA	Information security/ IT	Private	Reduction in viral threats	Crowdsourcing	Search, tracking, relationship analysis	Public
Phoebus	Big Data in Agriculture	France	Agriculture	Public	Improvement in agriculture worldwide	Data sharing partnerships	Search, text-mining, classification, collaborative analysis, and visualisation	Public
Lune	Traffic analysis Hub	UK	Human trafficking	Public	Elimination of human trafficking	Data sharing partnerships	Integration, single repository, tracking, Information sharing	Public
Sun	Indicavet	France/ UK	Food/ Antimicrobial consumption	Public	Reduction of antibiotics in food	Data sharing partnerships	Data tracking	Public

Notes: For more details on the 12 platforms cases, see the Web Appendix, Table A.

(marketplaces), extracting public insights from anonymised data sets (aggregators), or sharing data and valuable insights (disseminators) (See Fig. 4). Section 4.1 details these recurring patterns as they relate to developing platform archetypes.

4.1. Overview of platform archetypes

4.1.1. Data crawler

Crawlers are typically marketing agencies that process and accumulate publicly available data for commercial reasons. To do so, they convert publicly accessible unstructured (open) data (e.g., news, social media posts, research papers) to structured form (private good), thus facilitating market analysis, branding research, or ad targeting for a particular customer. These platforms leverage data from websites and social media (Mercury) and even initiate new social networks (Earth) to collect data in their area of interest. For instance, Earth provides stimuli for pet owners to post pictures of animals and interact. Crawlers apply data cleaning, integration, and classification algorithms (Mercury, Venus, and Earth) and use the analysis of the data set as private good. They capture value through standard contracts with the business-to-business (B2B) customer.

4.1.2. Data marketplace

Marketplaces unlock purchasing transactions between parties

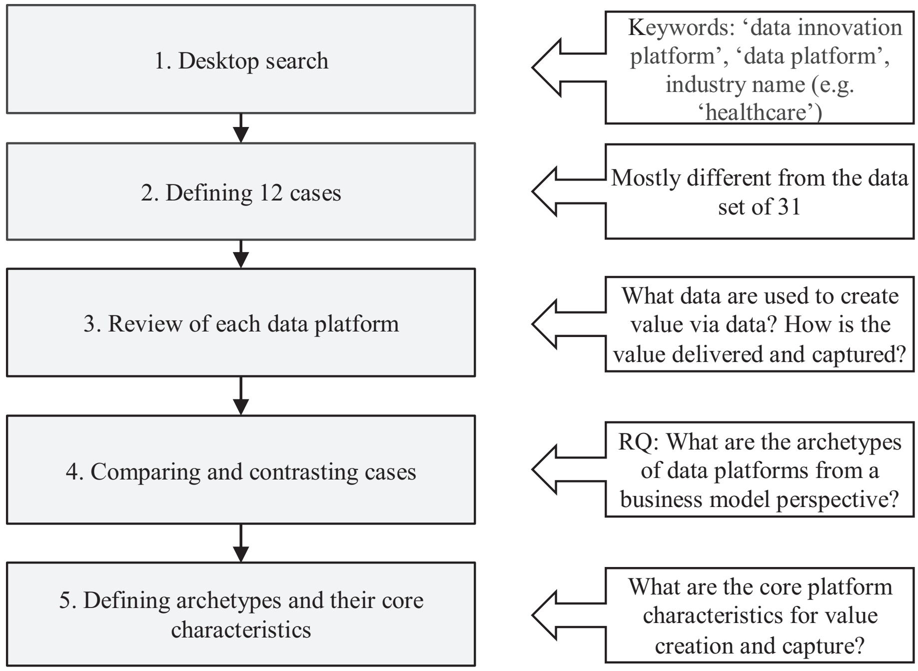

Fig. 3. A flow chart of data analysis.

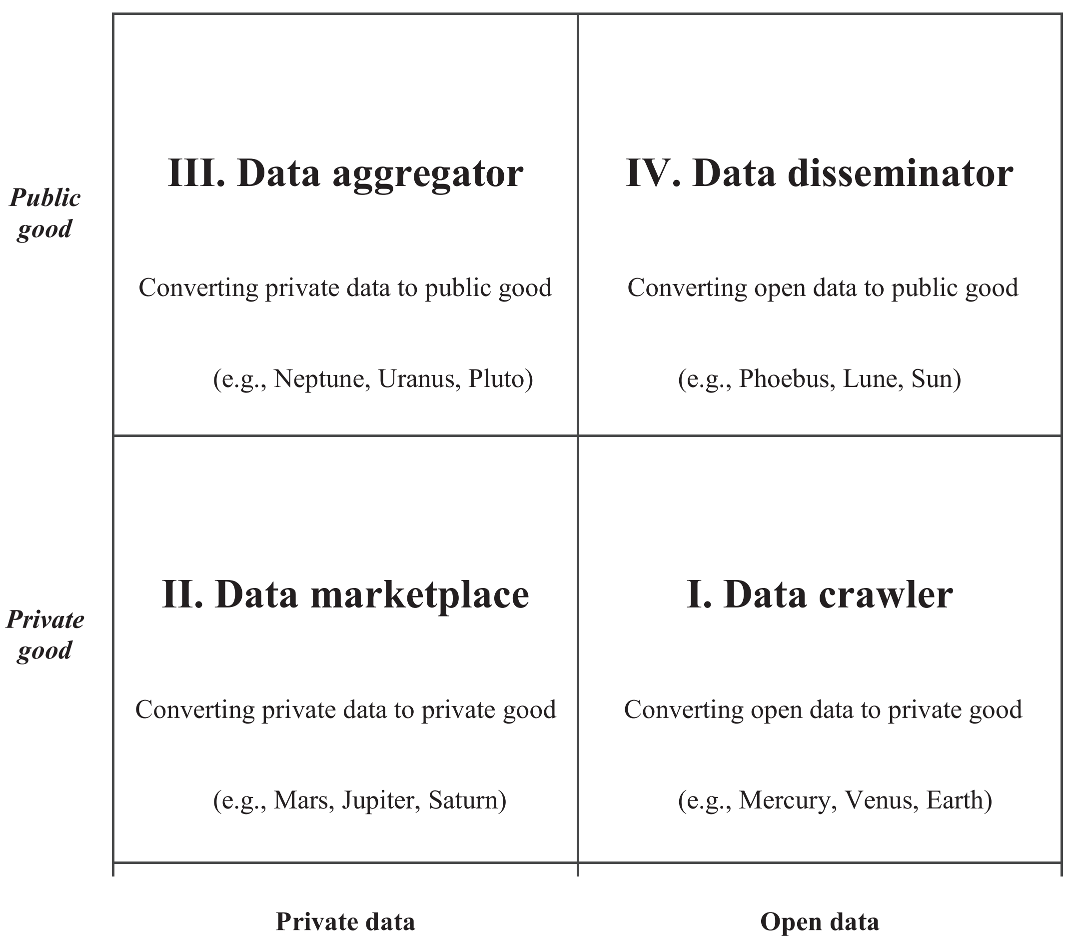

Fig. 4. Classification of business model archetypes using a 2 × 2 matrix.

interested in selling and buying private data (Mars, Jupiter, Saturn), thus providing private goods for the sellers. The value is created by enabling users to generate their own data sets (Mars, Saturn) or providing robust privacy protection mechanisms that make data available (Jupiter). Data-driven analytics, such as productivity improvements for livestock (Mars), may incentivise small-scale farmers to generate and make available data for the platform. While data-driven analytics improve production processes for these farmers, the company uses these data to optimise its livestock marketplace and attract sellers and buyers to commit to purchase transactions. Jupiter ensures the automated consent of patients to share their private medical data, which can be shared with physicians and healthcare organisations. Saturn provides a free DNA testing service to incentivise users to sell their data. Thus, by increasing the data supply and facilitating purchasing transactions with buyers, marketplaces benefit from sale commissions and create value from private data for a private good. They deliver value using platforms and capture value through a percentage commission in transactions.

4.1.3. Data aggregator

Aggregators accumulate private data by providing trustworthy data collection conditions and focusing on creating long-term benefits for society. They create value by uploading, integrating, and collaboratively analysing the anonymised data, such as electronic health records to enable the development of new treatments (Neptune), road condition data to improve safety and traffic flow (Uranus), and infected files to enable effective antiviral remedies (Pluto). The additional services allow better compliance with privacy standards (Saturn) and raise social capital (Pluto). Therefore, aggregators transform private data for the public good. They deliver value using platforms, and they capture value through standard contracts.

4.1.4. Data disseminator

Disseminators share massive data assets for the public good and provide a major impact via data sharing. The value is created by providing access to the data volumes and analytics to, for example, small agricultural farms, often in developing countries (Phoebus); intelligence reports and awareness campaigns (Lune); or sustainability projects (Sun). The disseminators use data-sharing partnerships (Phoebus, Lune, Sun) to populate the data set. By sharing data sets on a large scale and deriving insights where needed (Phoebus, Lune, and Sun), disseminators create value by converting public data into a public good. Disseminators deliver value using platforms and capture value through grants, donations, or subsidies.

4.2. Overview of archetypes

In this section, we compare these business model archetypes in detail. We apply two vertical dimensions in the framework, on which the archetypes share similar data privacy level, and two horizontal dimensions, on which the archetypes share similar application purpose.

The first dimension, which includes the crawlers (I) and marketplaces (II), applies to openly accessible (private data) for a private good. We differentiate these two orthogonal archetypes based on the size of the focal firm: Whereas crawlers deliver services to marketing departments of large customers, marketplaces attract individual data contributors and small and medium-sized enterprises (SMEs). Crawlers can collect data sets and sell them via marketplaces. As a result, they can create a new customer segment of SMEs that require data support for their operations. In turn, marketplaces can utilise social listening tools used by crawlers to collect data to increase the data supply.

The second dimension, which includes the marketplaces (II) and aggregators (III), pertains to the conversion of private data into commercial data for a societal need. We differentiate these two orthogonal archetypes based on longitudinal criteria – that is, whether they unlock sensitive data for immediate sale or accumulate these data for further analysis. We note that the long-term data-driven effect is often blended between private and public goods, as companies need to sustain themselves before helping others. For instance, marketplaces can increase the number of transactions by selling anonymised parts of a larger private data set of aggregators and generating private goods for data owners. At the same time, aggregators can benefit from the efficiency of marketplaces to convert data for value, thus covering their operational expenses and thereby sustaining their public good impact for the future.

The third dimension, which includes the aggregators (III) and disseminators (IV), is characterised by targeting public goods derived from private or open data. We differentiate these two orthogonal archetypes based on their demand-driven application character; that is, whereas aggregators develop a high-value service to target sustainable development goals, disseminators can make a difference by releasing terabytes of data on demand. Examples of such data sharing include data supply for recovering from ecological and humanitarian catastrophes (e.g., tsunamis, earthquakes, floods), finding lost people, preventing human trafficking, or furthering criminal investigations. Disseminators have also been called ‘data collaboratives’[3] in the literature because they provide critical, timely support during disasters by unlocking previously inaccessible data to solve societal problems (Verhulst et al., 2015). In contrast to aggregators, disseminators depend less on the crowdsourcing community due to their larger size. Aggregators can become disseminators by growing their data sets, gaining a reputation, and entering more data-sharing partnerships.

The fourth dimension, which includes disseminators (IV) and crawlers (I), involves the application of openly accessible data for public or private good. We differentiate these two orthogonal archetypes according to their impact leverage criteria: crawlers have a small niche impact, whereas disseminators have a large-scale impact. For crawlers, selling data-driven insights represents the main source of income, although they can also be co-funded by grants so that their open data collection strengths can be used for the public good. In this sense, crawlers can help disseminators increase their impact even further.

Table 2 introduces the core characteristics of archetypes as features that facilitate platform value creation and capture.[4] First, data characteristics describe the 4Vs of data in use: volume, veracity, velocity, and variety. Next, value creation describes the core data platform’s tools: accessing open data,[5] data generation, data accumulation, or sharing partnerships. Finally, the value capture column shows how the platform capitalises on data sharing (i.e., via reporting, providing high-value service provision, or receiving a percentage from transactions). By comparing these core characteristics, we can differentiate the properties of each business model archetype.

5. Discussion

Platform ecosystems are influenced by a mix of private and open data, which provides fruitful research ground. However, despite the obvious potential of open data, the recognition of data-sharing benefits and demands is yet to be fully understood. This section outlines the main theoretical contributions and managerial implications to show how data sharing can be considered for business model innovation.

5.1. Theoretical contribution

Our research responds to earlier studies noting the limited research on the purpose of business model innovation ‘beyond the profits of an individual firm’ (Snihur and Bocken, 2022) and the need to systematise profit-oriented incentives, which should lead to more data sharing and

[3] https://datacollaboratives.org/explorer.html?#data-pooling (accessed 06 Feb 2022).

[4] Value delivery occurs through the platform.

[5] That is, whether the platform collects openly accessible data from social networks, such as Facebook, Instagram, and Twitter.

Table 2
Platform archetypes: core data sharing characteristics for value creation and capture.

	Core data characteristics				Value creation				Value capture		
	Volume	Veracity	Velocity	Variety	Analysing open data	Selling data	Curating data	Releasing data on demand	Report sales	High-value service subscription	Transaction percentage
Crawler			x	x	x				x		
Marketplace		x		x		x					x
Aggregator	x	x					x	x		x	
Disseminator	x			x	x			x	x		

greater societal impact (Mehta et al., 2021). First, we answer our research question (What are the business model archetypes of platform ecosystems from a data-sharing perspective?) by proposing four archetypes of platforms. Building on these archetypes, we formulate the following propositions about data sharing on platforms and value realisation:

1. Data crawlers pool open data, create a large data set, clean and harmonise this data set, and derive data-driven insights around the product users.

Proposition 1. A platform owner that collects accessible industrial data can earn revenue by selling targeted reports.

2. Data marketplaces enable data contributors, such as individual users or SMEs, to commercialise their data (Vomfell et al., 2015) and strengthen data supply for data buyers (Deichmann et al., 2016; Spiekermann, 2019).

Proposition 2. A platform owner that matches industrial data sellers and buyers can earn revenue by orchestrating the data marketplace.

3. Data aggregators pool private data for the public good to enable collaborative analysis and open innovation (Chesbrough and Bogers, 2014).

Proposition 3. A platform owner that collects a private industrial data set can create a public good by developing a predictive analytics service.

4. Data disseminators provide large volumes of data from different sectors – including private companies and research institutions (Verhulst et al., 2015) – to help solve public problems (Stiglitz, 1999; Verhulst et al., 2015; Young and Verhulst, 2020), thereby leveraging data sharing to provide a substantive impact.

Proposition 4. A platform owner that enables industrial data partnerships can create a public good by releasing data on demand.

Second, considering that data platforms are often excluded from business model innovation research due to their lack of a profit-oriented nature (Fruhwirth et al., 2020), we posit the four archetypes as a mechanism for value realisation using data sharing. By disclosing data-related similarities and differences between business models in platform ecosystems, we show how the value from data sharing is created, delivered, and captured. By doing so, we extend the current knowledge of how data sharing impacts business model innovation (Janssen and Zuiderwijk, 2014; Kassen, 2018; Magalhaes and Roseira, 2020).

Third, by explaining how the archetypes function and evolve, we show how platforms can incentivise more ecosystem actors to unlock their private data for the purpose (Foss and Saebi, 2017; Teece, 2018). Moreover, we explain profit-oriented opportunities behind data sharing so that data owners can share their data more purposefully (Fruhwirth et al., 2020). Doing so facilitates a trade-off between data dissemination and commercialisation for ecosystem actors. The existing means to create trustworthy conditions or 'extract' more data from suppliers could potentially be applied for public good initiatives, especially when individuals do not trust data-sharing mechanisms or have not digitised their records, scenarios that represent an underexploited opportunity. If successful, data platforms can realise more value within a business or social context (Yoo et al., 2010) and for more societal actors (Magalhaes and Roseira, 2020; Zuiderwijk et al., 2014), thus positively impacting both business ecosystems and society (Foss and Saebi, 2017; Teece, 2018).

5.2. Managerial implications

The present study is based on real examples of platforms from 12 industries; thus, the implications for managers pertain to the application of findings to business model innovation. The archetypes can guide managers in developing more effective strategies to create more value from data. First, we encourage large organisations to consider the value of data sharing by using industrial relationships with their suppliers and customers. Second, we advise SMEs to take a broader view of the potential marketplaces for their generated data sets; these marketplaces may involve not just external but also existing data. For companies that cannot apply expensive data-driven analytical solutions, sharing data could enable actionable insights on the exchange of knowledge and capabilities, which would allow them to become more efficient by removing unnecessary waste from their business processes and thus enable growth opportunities. Third, we urge existing platform owners to consider applying the most appropriate archetypes according to this research to strengthen data collection, data analysis, and data-driven services for the private and public good.

6. Conclusion

In this paper, we investigate business models of platform ecosystems from a data-sharing perspective and derive four archetypes: crawlers, marketplaces, aggregators, and disseminators. They differ based on data privacy, core data characteristics, means to create and capture value, and public impact. As such, platform ecosystems can become data crawlers by adopting functionalities for scraping publicly available data and converting them to private goods, data marketplaces by enabling the trading functionality for buying and selling data sets for a private good, data aggregators by consolidating data and incentivising collaborative work on the data sets, or data disseminators by enabling data sharing partnerships and exploring the outcomes for societal or environmental impact (see Tables 1 and 2). We highlight the trade-off between data commercialisation and data dissemination in business model innovation of platforms – in particular, by securing a place for open data platforms in business model innovation and public good as an important value realisation outcome of business model innovation. Although data confidentiality, which is critical in cases such as human health data, can impede sharing, we note the large number of platform ecosystems in industries with greater data privacy and posit that unlocking the value of data there can provide the highest returns along both private and public good dimensions.

6.1. Research limitations

Whilst we contribute to the literature on business model innovation, we acknowledge several limitations in our research. First, we carried out the investigation using platform descriptions in English, which potentially limits the data selection, as the data platforms outside the English-speaking world were implicitly excluded from the sample. Second, we did not collect primary data about the selected 12 platforms; therefore, we cannot exclude the possibility that secondary data were misinterpreted.

6.2. Suggestions for future research

Researchers may follow up this research by considering languages other than English, thus expanding the scope of the analysed platforms and validating the archetypes discovered in this research. In addition, follow-up studies could investigate underrepresented B2B industries, (e.g., aerospace, automotive industries, food manufacturing). It would be particularly valuable to learn if these archetypes can transform themselves in time (i.e., from crawler to marketplace or from aggregator to disseminator) to realise more value from data sharing. An exciting opportunity for further research is to explore the role of actors in data innovation ecosystems. The present paper is focused on one actor – the digital platform owner – and the potential business models of platform ecosystems from a data-sharing perspective. Further research could focus on the roles of specific actors in platform ecosystems, such as data collaborators, and comparing/contrasting them with existing ecosystem actor roles (Jacobides et al., 2018; Iansiti and Levien, 2004). Studies in this vein could shed light on how to incentivise more actors to unlock private data for a shared societal purpose. In addition, follow-up research could include collecting primary data from innovation platforms, which could confirm or complement our study of archetypes. Finally, further quantitative studies could test the derived propositions using a broader sample of platforms.

CRediT authorship contribution statement

Nikolai Kazantsev: Conceptualization, Methodology, Data Curation, Formal analysis, Writing – Original Draft.
Nazrul Islam: Supervision, Conceptualization, Research Method & Validation, Review & Editing.
Jeremy Zwiegelaar: Supervision, Data Curation, Review & Editing.
Alan Brown: Supervision, Review & Editing.
Roger Maull: Supervision, Review & Editing.

Data availability

The authors are unable or have chosen not to specify which data has been used.

Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council (grant number EP/T022566/1). DIGIT Lab is a Next Stage Digital Economy Centre. We also acknowledge the role of Veterinary Health Innovation Engine (vHive) — a unique partnership between the University of Surrey and Zoetis Centre for Digital Innovation that bridges the technology readiness level gap between the world's leaders in animal health while leading in the development of transformational digital analytics tools to advance the well-being of animals for the benefit of society. In particular, we thank Dr. Theo Kanellos for their valuable input with regard to the digital economy and sustainability.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.techfore.2023.122515.

References

Adner, R., 2006. Match your innovation strategy to your innovation ecosystem. Harv. Bus. Rev. 84 (4), 98.
Adner, R., 2017. Ecosystem as structure: an actionable construct for strategy. J. Manag. 43 (1), 39–58.
Amit, R., Zott, C., 2001. Value creation in e-business. Strateg. Manag. J. 22 (6–7), 493–520. https://doi.org/10.1002/smj.187.
Baldwin, C.Y., Woodard, C.J., 2009. The architecture of platforms: a unified view. Platf. Mark. Innov. 32, 19–44.
Bock, M., Wiener, M., 2017. Towards a taxonomy of digital business models – conceptual dimensions and empirical illustrations, presentation at ICIS, Seoul, South Korea. December.
Bogers, M., Chesbrough, H., Moedas, C., 2018. Open innovation: research, practices, and policies. Calif. Manag. Rev. 60 (2), 5–16.
Bonina, C., Koskinen, K., Eaton, B., Gawer, A., 2021. Digital platforms for development: foundations and research agenda. Inf. Syst. J. 31 (6), 869–902. https://doi.org/10.1111/isj.12326.
Brown, A.W., 2021. Revitalizing your data-driven innovation strategy. In: Digital Economy Dispatch. https://alanwbrown.com/dispatches/dispatch-039/.
Candelon, F., Russo, M., di Carlo, R., Bedraoui, H.El, Feng, T., 2020. Simple governance for data ecosystems. December 2. In: BCG. https://www.bcg.com/publications/2020/simple-governance-rules-for-data-ecosystems.
Carnelley, P., Schwenk, H., Cattaneo, G., Micheletti, G., Osimo, D., 2013. Europe's data marketplaces – current status and future perspectives. European Data Market SMART 63. https://datalandscape.eu/data-driven-stories/europe%E2%80%99s-data-marketplaces-%E2%80%93-current-status-and-future-perspectives.
Chesbrough, H., Bogers, M., 2014. Explicating open innovation: clarifying an emerging paradigm for understanding innovation. In: New Frontiers in Open Innovation. Oxford University Press, Oxford, UK, pp. 3–28.
Chesbrough, H., Rosenbloom, R., 2002. The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies. Ind. Corp. Chang. 11 (3), 529–555.
Cozzolino, A., Corbo, L., Aversa, P., 2021. Digital platform-based ecosystems: the evolution of collaboration and competition between incumbent producers and entrant platforms. J. Bus. Res. 126, 385–400. https://doi.org/10.1016/j.jbusres.2020.12.058.
Cusumano, M.A., Gawer, A., Yoffie, D.B., 2019. The Business of Platforms: Strategy in the Age of Digital Competition, Innovation, and Power. Harper Business, New York.
Davies, T., Perini, F., 2016. Researching the emerging impacts of open data: revisiting the ODDC conceptual framework. J. Comm. Inform. 12 (2), 148–178. https://doi.org/10.15353/joci.v12i2.3246.
Deichmann, U., Goyal, A., Mishra, D., 2016. Will digital technologies transform agriculture in developing countries? Agric. Econ. 47 (S1), 21–33.
Demirkan, H., Spohrer, J., 2015. T-shaped innovators: identifying the right talent to support service innovation. Res.-Technol. Manag. 58 (5), 12–15.
Eisenhardt, K.M., 1989. Building theories from case study research. Acad. Manag. Rev. 14 (4), 532–550.
Eisenhardt, K.M., Graebner, M.E., 2007. Theory building from cases: opportunities and challenges. Acad. Manag. J. 50 (1), 25–32.
European Commission, 2016. Open science, open to the world – a vision for Europe. In: European Comission Report.
Foss, N.J., Saebi, T., 2017. Fifteen years of research on business model innovation. J. Manag. 43 (1), 200–227. https://doi.org/10.1177/0149206316675927.
Fruhwirth, M., Rachinger, M., Prlja, E., 2020. Discovering business models of data marketplaces. December 2. In: Hawaii International Conference on System Sciences.
Guba, E.G., Lincoln, Y.S., 1994. Competing paradigms in qualitative research. In: Denzin, N.K., Lincoln, Y.S. (Eds.), Handbook of Qualitative Research. SAGE Publications, Thousand Oaks, CA, pp. 105–117.
Hampton, S.E., Anderson, S.S., Bagby, S.C., Gries, C., Han, X., Hart, E.M., Jones, M.B., Lenhardt, W.C., MacDonald, A., Michener, W.K., Mudge, J., Pourmokhtarian, A., Schildhauer, M.P., Woo, K.H., Zimmerman, N., 2015. The Tao of open science for ecology. Ecosphere 6 (7), 120. https://doi.org/10.1890/ES14-00402.1.
Hartmann, P.M., Zaki, M., Feldmann, N., Neely, A., 2016. Capturing value from big data – a taxonomy of data-driven business models used by start-up firms. Int. J. Oper. Prod. Manag. 36 (10), 1382–1406. https://doi.org/10.1108/ijopm-02-2014-0098.
Hein, A., Schreieck, M., Riasanow, T., Setzke, D.S., Wiesche, M., Böhm, M., Krcmar, H., 2020. Digital platform ecosystems. Electron. Mark. 30 (1), 87–98.
Huikkola, T., Rabetino, R., Kohtamäki, M., Gebauer, H., 2020. Firm boundaries in servitization: interplay and repositioning practices. Ind. Mark. Manag. 90, 90–105.
Huyer, E., van Knippenberg, L., 2020. The economic impact of open data: opportunities for value creation in Europe. Eur. Comm.
Iansiti, M., Levien, R., 2004. The keystone advantage: What the new dynamics of business ecosystems mean for strategy, innovation, and sustainability. Harvard Business Press, Boston.

Immonen, A., Palviainen, M., Ovaska, E., 2014. Requirements of an open data based business ecosystem. IEEE Access 2, 88–103.
Jacobides, M.G., Cennamo, C., Gawer, A., 2018. Towards a theory of ecosystems. Strat. Manag. J. 39 (8), 2255–2276.
Janssen, M., Zuiderwijk, A., 2014. Infomediary business models for connecting open data providers and users. Soc. Sci. Comp. Rev. 32 (5), 694–711. https://doi.org/10.1177/0894439314525902.
Kassen, M., 2018. Open data and its institutional ecosystems: a comparative cross-jurisdictional analysis of open data platforms. Can. Public Adm. 61 (1), 109–129. https://doi.org/10.1111/capa.12251.
Kundra, V., 2012. Digital fuel of the 21st century: Innovation through open data and the network effect. Joan Shorenstein Center on the Press, Politics and Public Policy, Boston.
Linde, L., Sjödin, D., Parida, V., Wincent, J., 2021. Dynamic capabilities for ecosystem orchestration: a capability-based framework for smart city innovation initiatives. Technol. Forecast. Soc. Change 166, 120614.
Magalhaes, G., Roseira, C., 2020. Open government data and the private sector: an empirical view on business models and value creation. Gov. Inform. Q. 37 (3), 101248 https://doi.org/10.1016/j.giq.2017.08.004.
Mamonov, S., Triantoro, T.M., 2018. The strategic value of data resources in emergent industries. Int. J. Inform. Manag. 39, 146–155. https://doi.org/10.1016/j.ijinfomgt.2017.12.004.
Manyika, J., Chui, M., Groves, P., Farrell, D., Kuiken, S.Van, Doshi, E.A., 2013. In: Open data: unlocking innovation and performance with liquid information, 21. McKinsey Global Institute, p. 116.
Massa, L., Tucci, C.L., Afuah, A., 2017. A critical assessment of business model research. Acad. Manag. Ann. 11 (1), 73–104. https://doi.org/10.5465/annals.2014.0072.
Mehta, S., Dawande, M., Mookerjee, V., 2021. Can data cooperatives sustain themselves?. In: London School of Economics Blog https://blogs.lse.ac.uk/businessreview/2021/08/02/can-data-cooperatives-sustain-themselves/.
Miles, M.B., Huberman, A.M., 1994. Qualitative data analysis: An expanded sourcebook. SAGE Publications, Thousand Oaks, CA.
Ng, I.C., Nudurupati, S.S., Tasker, P., 2010. Value co-creation in the delivery of outcome-based contracts for business-to-business service. In: AIM Working Research Paper. https://www.bl.uk/collection-items/value-cocreation-in-the-delivery-of-outcome based-contracts-for-business-to-business-services.
Osterwalder, A., Pigneur, Y., 2010. Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers, Vol. 1. John Wiley & Sons, New York.
Otto, B., Aier, S., 2013. Business models in the data economy: a case study from the business partner data domain. In: Wirtschaftsinformatik Proceedings. https://aisel.aisnet.org/wi2013/30.
Parida, V., Burström, T., Visnjic, I., Wincent, J., 2019. Orchestrating industrial ecosystem in circular economy: a two-stage transformation model for large manufacturing companies. J. Bus. Res. 101, 715–725. https://doi.org/10.1016/j.jbusres.2019.01.006.
Reim, W., Parida, V., Örtqvist, D., 2015. Product–service systems (PSS) business models and tactics – a systematic literature review. J. Clean. Prod. 97, 61–75. https://doi.org/10.1016/j.jclepro.2014.07.003.
Salzberger, T., Sinkovics, R.R., Schlegelmilch, B.B., 1999. Data equivalence in cross-cultural research: a comparison of classical test theory and latent trait theory based approaches. Australas. Mark. J. 7 (2), 23–38.
Samuelson, P.A., 1954. The pure theory of public expenditure. Rev. Econ. Stat. 36 (4), 387–389. https://doi.org/10.2307/1925895.
Sinkovics, R.R., Penz, E., Ghauri, P.N., 2008. Enhancing the trustworthiness of qualitative research in international business. Manag. Int. Rev. 48 (6), 689–714.
Snihur, Y., Bocken, N., 2022. A call for action: the impact of business model innovation on business ecosystems, society and planet. Long Range Plan. 55 (6), 102182 https://doi.org/10.1016/j.lrp.2022.102182.
Spiekermann, M., 2019. Data marketplaces: trends and monetisation of data goods. Intereconomics 54 (4), 208–216.
Stahl, F., Schomm, F., Vossen, G., 2014. The data marketplace survey revisited. In: ERCIS working paper. https://www.econstor.eu/handle/10419/94187.
Stiglitz, J.E., 1999. Knowledge as a global public good. In: Kaul, I., Grunberg, I., Stern, M. (Eds.), Global Public Goods: International Cooperation in the 21st Century. Oxford University Press, Oxford, pp. 308–325.
Teece, D.J., 2010. Business models, business strategy and innovation. Long Range Plan. 43 (2–3), 172–194. https://doi.org/10.1016/j.lrp.2009.07.003.
Teece, D.J., 2018. Business models and dynamic capabilities. Long Range Plan. 51 (1), 40–49. https://doi.org/10.1016/j.lrp.2017.06.007.
Teece, D.J., Linden, G., 2017. Business models, value capture, and the digital enterprise. J. Organ. Des. 6 (1) https://doi.org/10.1186/s41469-017-0018-x.
Thomas, L.D., Leiponen, A., 2016. Big data commercialization. IEEE Eng. Manag. Rev. 44 (2), 74–90. https://doi.org/10.1109/emr.2016.2568798.
Tiwana, A., Konsynski, B., Bush, A.A., 2010. Research commentary – platform evolution: coevolution of platform architecture, governance, and environmental dynamics. Inf. Syst. Res. 21 (4), 675–687. https://doi.org/10.1287/isre.1100.0323.
Verhulst, S., Young, A., Srinivasan, P., 2015. An Introduction to Data Collaboratives, GovLab Report. https://datacollaboratives.org/static/files/data-collaboratives-intro.pdf.
Vomfell, L., Stahl, F., Schomm, F., Vossen, G., 2015. A classification framework for data marketplaces (No. 23). Ercis working paper.
Welch, C., Piekkari, R., 2017. How should we (not) judge the 'quality'of qualitative research? A re-assessment of current evaluative criteria in international business. J. World Bus. 52 (5), 714–725.
Wixom, B.H., Ross, J.W., 2017. How to monetize your data. MIT Sloan Management Review 58 (3). https://sloanreview.mit.edu/article/how-to-monetize-your-data/#:~:text=Companies%20can%20take%20three%20approaches,to%20new%20and%20existing%20markets.
Yin, R.K., 1993. Case Study Research Design and Methods Applied. In: Social Research Methods Series, 5. SAGE Publications, Thousand Oaks, CA.
Yin, R.K., 2009. Case Study Research: Design and Methods, Vol. 5. SAGE Publications, Thousand Oaks, CA.
Yoo, Y., Henfridsson, O., Lyytinen, K., 2010. Research commentary – the new organizing logic of digital innovation: an agenda for information systems research. Inf. Syst. Res. 21 (4), 724–735. https://doi.org/10.1287/isre.1100.0322.
Young, A., Verhulst, S.G., 2020. Data collaboratives. In: Harris, P., Bitonti, A., Fleisher, C., Skorkjær, B.A. (Eds.). The Palgrave Encyclopedia of Interest Groups, Lobbying and Public Affairs, Cham: Palgrave Macmillan.
Zuiderwijk, A., Loukis, E., Alexopoulos, C., Janssen, M., Jeffery, K., 2014. Elements for the development of an open data marketplace. In: Presentation at Conference for E-democracy and Open Government, pp. 309–322.

Nikolai Kazantsev is a postdoctoral research associate in the Elastic Manufacturing Project, Institute for Manufacturing, University of Cambridge, UK. He worked before in University of Exeter, UK and in Higher School of Economics Moscow. His research interests are related to data-driven business model innovation and the orchestration of nascent ecosystems in manufacturing. Earlier Nikolai has completed PhD in Business and Management at the University of Manchester (Alliance Manchester Business School) and an MSc degree in Information Systems at the University of Muenster (Germany).

Nazrul Islam is Chair Professor of Business & Director of Research Degrees at the Royal Docks School of Business and Law, University of East London, UK. He holds a PhD in innovation management. His research interest focuses on interdisciplinary fields: the management of technology; technological transformation; the emergence and growth of disruptive and digital technology-based innovation; and SMEs business sustainability. His research was published in the leading international journals, and he has complemented his peer reviewed journal efforts with three books. Prof Islam's research received awards including the 'Brad Hosler Award for Outstanding Paper' from USA; and the 'Pratt & Whitney Canada Best Paper Award' from Canada. Prof Islam serves on the board of directors for *Business and Applied Sciences Academy of North America.* He is an Associate Editor for *Technological Forecasting & Social Change*, Department Editor for *IEEE Transactions on Engineering Management*, and Editor-in-Chief of *International Journal of Technology Intelligence and Planning.*

Jeremy Zwiegelaar is currently the Subject coordinator for Entrepreneurship at Oxford Brookes University, UK. He has a PhD in Management (Massey University NZ), MA Hons 1st class Psychology (University of Auckland, NZ) and a Graduate Diploma in Teaching (University of Auckland, NZ). He is researching in the area of nascent entrepreneurship with a focus on new venture performance. He studies entrepreneurs at the beginning of their business venture stage and consider aspects of performance, success and failure outcomes, and factors affecting or explaining this phenomenon. His work has appeared in *Journal of Small Business Management*, *R and D Management*, *Strategic Change* among others.

Alan Brown is Professor in Digital Economy at the University of Exeter's Business School, UK. Alan's research is focused on agile approaches to business transformation, and the relationship between technology innovation and business innovation in today's rapidly-evolving digital economy. After receiving a PhD in Computational Science at the University of Newcastle-upon-Tyne, Alan spent almost 20 years in the USA in commercial high-tech companies leading R&D teams, building leading-edge solutions, and driving innovation in software product delivery. Alan co-founded the Surrey Centre for the Digital Economy (CoDE) at the University of Surrey where he led research activities in 4 EPSRC-funded research projects. His research provides practical insights into the effects of digital disruption in business and society, and to deliver tools that offer measurable impact to drive strategic and operational advantages in the emerging digital economy.

Roger Maull is Professor of Management Systems at the University of Exeter's Business School, UK and Academic Director of the Initiative for the Digital Economy at Exeter (INDEX), which is based on London's South Bank. Roger's initial research interest focused on Operations Management, specifically in the services sector and he has continued to publish widely on that topic. Since 2009 his research has moved to focus on the impact of the Digital Economy on individuals, organisations and more generally on society. He is (as of April 2021) an investigator on eight UKRI projects.