Journal of Business Research 168 (2023) 114244

Contents lists available at ScienceDirect

Journal of Business Research

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

A pathway to bypassing market entry barriers from data network effects: A case study of a start-up's use of machine learning

Darek M. Haftor [a,b,*], Ricardo Costa-Climent [a,b], Samuel Ribeiro Navarrete [b,c]

[a] Uppsala University, Box 513, SE-75120 Uppsala, Sweden
[b] University of Economics and Human Sciences, Warsaw, Poland
[c] ESICS University Madrid, Spain

ARTICLE INFO

Keywords:
Artificial intelligence
Machine learning
Market entry
Legitimacy
SMEs
Stakeholders

ABSTRACT

Highly valued firms exploit machine learning to activate data network effects. Data is gathered and analyzed to generate predictions and recommendations. This loop locks in existing service users and locks out potential competitors, thus creating a sizeable entry barrier, particularly for small and medium-sized (SME) enterprises. The literature does not describe the possible pathways to enter markets protected by incumbents' data network effects. This study examines an SME that successfully entered such a market. A key finding is that, for successful market entry, an SME can focus on different stakeholders from those that are targeted by incumbents, provided such stakeholders can legitimize the SME's use of user data generated by incumbents.

1. Introduction

The prediction of demand is one of the most important bases for enterprise decision-making and planning, which is important to reduce risks and avoid blindness.

The smartphone application *Waze* provides turn-by-turn navigation recommendations that are sensitive to existing and emerging traffic situations. This service offers recommendations for the fastest routes, which might not be the shortest. The app has a large community of users, who are continuously and automatically sharing their spatial and temporal travel data and who also provide data on situations such as traffic jams and accidents. That data, together with historical data and other kinds of data such as weather conditions, enables Waze to use various machine learning techniques (Agrawal, Gans, & Goldfarb, 2018; Collins, Dennehy, Conboy, & Mikalef, 2021; Finlay, 2017) to identify patterns. These patterns are then used to produce predictions and offer recommendations for the most efficient travel routes. Demand forecasting is one of the most important bases for business decision-making and planning, which is important to reduce risk and prevent blindness (Lin, Lin, & Wang, 2022). The superior accuracy of recommendations by Waze has made it so popular that it has retained its current users while attracting many new users. This growth in its user base has enabled Waze to generate even more travel data, which is used to update the accuracy of its predictions and thus its route recommendations. This self-reinforcing positive learning loop is an example of a data network effect. Network integration is an essential determinant of business innovation, but the mechanism of its impact on innovation performance remains unclear (Boxu, Xingguang, & Kou, 2022). Such a data network effect must be distinguished from a conventional network effect, which depends on the size of the network (Gregory, Henfridsson, Kaganer, & Kyriakou, 2021). Once a firm has activated data network effects, its services will improve thanks to the user data it collects. Thus, it becomes difficult, if not impossible, for competitors to enter the market or remain relevant. As if to reinforce this idea, Google acquired Waze for nearly 1 billion USD in 2013 after Waze had been in operation for five years (Kerr, 2013). The activation of data network effects has contributed to the growth of several other well-known firms at a pace and magnitude never seen before (Parker, Van Alstyne, & Choudary, 2016). Examples include Google's search engine, Amazon's recommendations of products to customers, and Facebook's matching of ads with members' profiles.

When a pioneering firm establishes user services and gradually activates network effects, the users of those services are reluctant to migrate to alternative service providers because of the excellent service quality they receive based on the superior predictions enabled by large, unique sets of data. This lock-in of users is associated with the pioneering firm's acquisition of a large set of user data that new entrants lack and will be unlikely to generate themselves as long as users remain locked in to the pioneering firm's services. Well-activated data network

* Corresponding author at: Uppsala University, Box 513, SE-75120 Uppsala, Sweden.
E-mail address: darek.haftor@im.uu.se (D.M. Haftor).

https://doi.org/10.1016/j.jbusres.2023.114244
Received 24 December 2022; Received in revised form 19 August 2023; Accepted 24 August 2023

effects therefore serve as a strong market entry barrier. Furthermore successful leadership in digital transformations are supported by the global attitude and mentality of the leader (Tagscherer & Carbon, 2023).

Bypassing such an entry barrier is even more challenging for a *small and medium-sized enterprise* (SME). In general, SMEs have become prominent in the sustainable development agenda due to their contribution to promoting economic growth, employment generation and poverty reduction (Karmaker, Al Aziz, Palit, & Bari, 2022). SME managers may be wary of innovations due to the financial costs of adopting innovative technologies (Abbasi, Alam, Du, & Huynh, 2021) and the associated uncertainty. Although typical SMEs may have the advantage of an innovation capability unhindered by massive path dependency (Thietart, 2016), they often lack a large enough customer base and positive cash flows, as well as the financial, human, and other resources necessary to compete head-to-head with incumbents (Macpherson, Jones, Zhang, & Wilson, 2003; Supyuenyong, Islam, & Kulkarni, 2009; Wong & Aspinwall, 2005). Empirical evidence shows that value proposition innovation drives business performance, while value creation and value capture innovation serve as two key vectors in the mechanism (Guo, Guo, & Ma, 2022). Because SMEs have relatively limited resources, their survival depends on their adoption of new technologies (Abbasi et al., 2021). Current theorizations of SME market entry indicate several ways to bypass entry barriers: (a) by pursuing a broader product and market scope that gives rise to superior cost efficiencies, (b) by differentiating products further, thus attracting a customer base, (c) by targeting a market niche where customers are underserved by incumbents, thus expanding the customer base, or (d) by establishing an alliance with an incumbent (Acs, 2006; Arendt, 2008; Demsetz, 1982; Pehrsson, 2009). However, all these forms of bypassing entry barriers are invalidated by incumbents that have already activated strong data network effects.

Theorization of the activation of data network effects is recent (Gregory et al., 2021), so there are no studies of how to bypass such entry barriers. At the same time, key managerial questions arise. For example, is it possible for a start-up to enter a marketplace that is dominated by one or two incumbents that have activated strong data network effects? The market for mobile phone operating systems offers just such a case, with Apple's iOS and Google's Android leading the market. Hence, the following research question is addressed: ***What pathway can SMEs follow to enter a market with entry barriers that arise from incumbents' data network effects?*** An entry pathway refers to the sequence of decisions and subsequent actions that enable a firm to enter a market (Terlaak & Kim, 2021). This study's aim is therefore to identify a start-up's successful pathway into a market that incumbents dominate thanks to data network effects.

This question must be answered for at least two reasons. The first is the well-known market dominance of some firms that have activated data network effects (e.g., Amazon, Apple, Google, and Facebook), which regulators wish to deal with to re-establish competition and drive customer value creation (Smyth, 2019). The second is that, despite substantial potential benefits and an annual investment of some 55 billion dollars over the period 2015 to 2020 (Statista, 2021), there has been a slower adoption rate of artificial intelligence (AI) technologies among SMEs (Horváth & Szabó, 2019). In Europe, "75 % of European Union (EU) companies will be using cloud/artificial intelligence (AI)/ big data, and more than 90 % of small and medium-sized enterprises (SMEs) should reach at least a basic level of digital intensity by 2030" (Chen, Sun, & Chen, 2022: 1). However, SMEs account for the majority of businesses globally (World Bank, 2020), and they have a latent opportunity to make innovative use of AI technology (Sun, Zhao, & Sun, 2020).

To answer the present research question and thereby contribute to an important body of knowledge, this paper reports the results of a longitudinal exploratory study (Eisenhardt & Graebner, 2007) of a start-up that provides digital disease management services for hemophiliacs. The start-up used a novel approach to bypass the entry barriers established by large incumbent firms and thereby enter its target market. Through abductive reasoning (Tavory & Timmermans, 2014) that juxtaposes this case with the recently proposed theory of data network effects (Gregory et al., 2021; Kushwaha, Kar, & Dwivedi, 2021), this paper identifies a pathway to market entry. This pathway to market entry relies on legitimacy from stakeholders (Freeman, 1984; Freeman, Harrison, & Wicks, 2007) that have the authority to grant access to user data, including data generated by digital services provided by incumbents (Weerakkody, Irani, Kapoor, Sivarajah, & Dwivedi, 2017). In particular, the value proposition describes the organization's production department, as well as customer segments and relationships with a wide variety of stakeholders (Trapp, Kanbach, & Kraus, 2022). The results make several contributions to theory development. The recently proposed theory of data network effects benefits from two developments. First, the theory benefits from expanding the notion of a ***single human user*** of services based on machine learning techniques to the notion of ***multiple, interdependent users***, which may be individuals or organizations. Second, user value may be the result of the co-evolutionary unfolding of events, accounting for the temporality and dynamics of data network effects. Third, this study makes a key contribution to the literature on the market entry of SMEs. Specifically, it describes a novel way to bypass incumbents' entry barriers created by activated data network effects. These results have clear implications for firms' data management, business strategy, and business model design, as well as anti-trust policymaking.

The paper continues with a summary of the recently proposed theory of data network effects, which is assumed here to be a source of abductive reasoning. Thereafter, the research methods are detailed, and an account of the investigated firm is provided. This account is followed by a theoretical generalization based on learning from the case. The paper ends with a discussion of the findings and conclusions regarding the contribution of this study, as well as its practical implications and limitations.

2. Theoretical background

The recently proposed theory of data network effects (Gregory et al., 2021) serves as the primary lens for the empirical investigation to respond to the present research question. First, however, that theory must be contextualized within the more established theory of network effects. Strategic networks, understood as configurations of interactor ties that are central for value creation (Burt, 1992), include alliances, joint ventures, and long-term buyer–supplier partnerships (Gulati, Nohria, & Zaheer, 2000). A key underlying notion of such networks is that the locus of value creation is within a given strategic network, not necessarily within a single actor such as a firm (Doz & Hamel, 1998; Giusti, Alberti, & Belfanti, 2020; Gulati, 1999; Hilmersson, & Hilmersson, 2021). Recent advances in the use of digital technology have enabled interactions between actors that were not possible before (Adamides & Karacapilidis, 2020; Dutta & Segev, 1999; Rehm, Goel, & Junglas, 2017). These technology uses have enabled the activation of network effects as a distinct source of value creation from actor networks.

2.1. Direct and indirect network effects

Scholars and practitioners have known for some time that firms, particularly technology firms, can design their business models to activate network effects. This notion means that an offering's value comes from both the inherent characteristics of the offering (goods or services), which generate value through actual use, and the network of users that use the offering (Economides, 1996; Farrell & Saloner, 1986; Katz & Shapiro, 1985; Liebowitz & Margolis, 1994; Sheremata, 2004). A classic illustration of these effects is the telephone network, where a single telephone network user will generate no network effects and probably no value at all. In contrast, two or more users will generate network effects because the network's value increases with the number of users that the network can link together. That is, the more users of a telephone network there are, the more value it offers its users. The foundational business model of the social media firm Facebook is similar. The more members it has, the more links to other users it can offer and thereby the greater the value it provides (Eisenmann, Parker, & Van Alstyne, 2011).

This direct, or single-sided, network effect contrasts with indirect, or multi-sided, network effects. Indirect network effects occur when value comes from the volume of interactions between two or more actors involved in the generation of an offering, typically a producer and a consumer. The well-known mobility-as-a-service firm Uber illustrates this notion. The more riders there are in its network, the more drivers will be attracted to offer transport, which will produce income. Moreover, the more drivers there are in the network, the more riders will be attracted because pick-up times will be shorter and fees will fall as drivers' idle time decreases (Boudreau & Jeppesen, 2015; Hagiu, 2015; Rochet & Tirole, 2003). Some of the world's most valuable firms have created wealth by using just such a network effect, at a magnitude and pace never seen before. Notable examples include Alphabet (Google), Amazon, Apple, and Facebook (Parker et al., 2016).

2.2. Data network effects

Enabled by recent advances in digital technologies and their widespread adoption, firms have gathered large sets of data and have used machine learning technology to analyze those data sets (Alter, 2021; Duan, Edwards, & Dwivedi, 2019). This analysis enables a new kind of network effect, namely data network effects (Gregory et al., 2021). Whereas the core mechanism of direct and indirect network effects is the *size of the network* (Afuah, 2013), the core mechanism of data network effects is the *scale of learning* (Gregory et al., 2021).

Machine learning is understood here as the use of computer algorithms (statistical and mathematical optimization) to improve an offering, where data analysis enables such algorithms to produce accurate predictions. The recent theory of data network effects accounts for value creation from the use of machine learning in organizations that succeed in harnessing data network effects (Gregory et al., 2021). This theory holds that a provider of an offering develops a data network effect when there is a loop of recurring activities. Specifically, a provider of an offering collects user data and other relevant data. This data is then analyzed to detect patterns that enable predictions related to the offering. These predictions in turn serve as recommendations for user decisions. The actual decisions and their outcomes are monitored so that new data is collected. This new data helps update the previously identified data patterns and thereby improve the quality of recommendations. This loop then continues to iterate.

The scale of such data-driven learning and improvements creates increased value for users. Having more data enables learning, which gives rise to better predictions and recommendations, which enhances value creation, which drives up the use of the offering, which then enhances learning, and so on (Agrawal et al., 2018; Samuel, 1959). This loop is again illustrated by the route recommendation service Waze (Kela-Madar, 2021). Crucially, Waze collects location-dependent information, such as map data, travel times, and traffic information from users, as well as users' reports of accidents and traffic jams. Based on historical and real-time data on traffic, weather, and other factors, machine learning techniques generate recommendations for routes and offer real-time updates. This setup means that the fastest route may not always be the shortest one at a given time and place. The success of Waze has attracted many users, enabling it to generate large volumes of data. These volumes of data then improve the quality of its route recommendations. Such improvements attract even more users, who generate additional usage data, further improving the quality of predictions, and so on. Accordingly, Waze reflects a provider's capability "to learn from data to continuously improve its products or services for each user" (Gregory et al., 2021: 538). This process drives a user's utility, understood as "a function of the scale of data-driven learning and improvements" made with machine learning (Gregory et al., 2021: 536).

A core question answered by the theory of data network effects is the following: What factors co-condition user value creation from the use of machine learning technology? The answer is that data network effects have a direct positive relationship with a provider's machine learning capability and users' perceived value from the use of the provider's offering (Gregory et al., 2021). The theory further proposes that this relationship is moderated by the provider's legitimation, data stewardship, and user-centric design. Based on the work of Gregory et al. (2021), these co-conditioning factors are briefly summarized because several of them are crucial to understand the present empirical investigation and the subsequent response to the research question at hand.

2.3. Factors that co-condition users' perceived value

Machine learning capability is the core driver of data network effects, which derive from the ability to learn from data and thereby continuously improve offerings for each user. To generate user value, the identified data patterns enable predictions that are the core mechanism for user value creation. Such predictions enable the forecasting of occurrences to produce recommendations for a given decision. Examples of such decisions include film choices, medicine dosages, and credit worthiness. To this end, the *speed* and *accuracy* of predictions are regarded by the theory of data network effects as central for users' perceived value. If it takes an hour to recommend a 15-minute route, then users' perceived value will probably be lower than if the time to generate such a recommendation is only a second. Similarly, if the recommendation says that the route requires 15 min when it actually requires two hours, users' perceived value will probably also be low.

The next two co-conditioning factors concern *data quality* and *data quantity*. To produce accurate predictions, accurate patterns must be detected in available data sets. This detection of accurate patters in turn requires the right amount and quality of data. The quality and quantity of data depends on the specific application, the algorithmic techniques, and the requirements for user value creation. More generally, however, situations with an insufficient amount and quality of data about a range of cases of the same class must be avoided because they may lead to misinterpretations of new cases that algorithms encounter when used (Agrawal et al., 2018).

The next two co-conditioning factors concern user-centric design, understood as a user's *performance expectancy* and *effort expectancy* (Dwivedi, Rana, Janssen, Lal, Williams, & Clement, 2017). Because the quality of predictions is highly conditioned by the quality and quantity of the data sets analyzed for pattern recognition, it is crucial to incentivize potential users to use the offering, which is what generates new data. With this aim, *performance expectancy* is understood as the extent to which potential users assume that using the offering will enable them to complete the task at hand. *Effort expectancy*, on the other hand, is the extent to which potential users assume that using the offering will not require substantial effort. Both these factors shape potential users' intentions to use an offering and thereby to create value through usage and generate data for future predictions (Dwivedi, Rana, Jeyaraj, Clement, & Williams, 2019).

The final two factors concern the offering's legitimacy. The provider of an offering needs to secure the positive legitimacy of key stakeholders in order to ensure positive perceptions and therefore usage of an offering. Legitimacy is defined here as an actor's perceptions and assumptions that the decisions and subsequent actions of another actor are appropriate and desirable within a given social sphere of norms (Suchman, 1995). Therefore, if the navigation software Waze collected data on daily travel routes and provided this data to the government intelligence agency without users' acceptance, its users would most probably cease to use Waze. This scenario reflects the *personal data use* factor of an offering's legitimacy, which is understood as the provider's legitimate collection and use of personal data without violating the formal and informal norms for such data use. The second factor of an offering's legitimacy is *prediction explainability*, which concerns stakeholders' understanding of an offering with regard to why a recommendation is generated the way it is. A detailed understanding of the underlying mechanism used to generate a recommendation is often not possible due to its complexity (Mayenberger, 2019). However, the general underlying principles of how an offering generates predictions can be disclosed to ensure stakeholders' trust and understanding. For instance, if Waze recommended a route that was longer than the shortest available one but explained that the reason for the recommendation was to enable emergency vehicles to reach their destination faster, then users may find the recommendation appropriate.

Combining all the co-conditioning factors listed here, the theory of data network effects explains that a *users' perceived value* comes from the machine learning capability of a provider with a level of *prediction speed* and *accuracy* that is suitable for the given recommendation. This positive relationship is strengthened by higher *data quality* and *data quantity*, higher *performance expectancy*, lower *effort expectancy*, and more legitimate *personal data use* and *prediction explainability*. The theory of data network effects offers a unique account of value creation from the use of machine learning technology (Gregory et al., 2021), in contrast with the literature that accounts for value creation from the use of digital technologies in general (Bharadwaj, 2000; Brynjolfsson & Saunders, 2009; Chae, Koh, & Prybutok, 2014). The latter typically black-boxes technology by disregarding its unique characteristics and capabilities (Dwivedi et al., 2015; Orlikowski & Iacono, 2001), whereas the theory of data network effects explicitly accounts for a particular kind of digital technology use, namely machine learning. Despite its merits, this theory was recently criticized for its disregard of value *appropriation* from activated data network effects based on the use of machine learning (Clough & Wu, 2022; Gregory, Henfridsson, Kaganer, & Kyriakou, 2022). Although that criticism is valid, any theory has its limitations, and the theory of data network effects focuses on perceived value creation, not value appropriation.

2.4. Implications for market entry

A unique feature of data network effects is that value creation comes from many actors' use of an offering. The accuracy of the recommended travel routes provided by Waze depends on data generated by multiple users, both past and present. Hence, the more users there are, the better the recommendations will be, and the higher users' perceived value will become. By extension, this argument implies that if provider A is first to enter a market with a specific offering that activates data network effects (e.g., recommendations for travel routes, music, or films), then provider A will attract users with its novel offering, assuming that it is relevant to users' needs. As users start to use the new offering, they generate unique data that is crucial for the further improvement of the predictions that are central to that offering. This quality improvement is likely to attract new users and retain existing ones, which generates additional data and thereby enhances the quality of predictions. When provider B wishes to enter the same market with a similar offering, it is at a major disadvantage because it has access to neither historical usage data nor a current user base with real-time usage data. Therefore, the first mover (i.e., provider A) has established a crucial entry barrier for any imitator to enter the market.

A recent example is in the streaming music industry, where Spotify's early entry and rapid growth activated all three kinds of network effects (direct, indirect, and data network effects). Machine learning finds patterns regarding users' music preferences in a given situation (e.g., time, place, and personal profile). Thus, Spotify can recommend a playlist that is more accurate than those recommended by late-entry actors such as Apple Music (Saura, 2021; Towse, 2020). In this sense, Spotify has built an entry barrier to imitators from its activation of data network effects, even though Apple has a significantly larger user base for its total service provision. In the case of large and highly diversified firms, such as Alphabet (Google), Amazon, and Apple, late entry with inferior service provisions can be subsidized by other business lines aimed at providing complementary offerings (Amit & Zott, 2001). Such a strategy aims to make users of the first-entry provider migrate to late-entry providers, attracted by a portfolio of multiple complementary offerings that the first mover does not have. Depending on the specific case, such a strategy may or may not succeed (Watson, Weaven, Perkins, Sardana, & Palmatier, 2018). Because SMEs are typically single-business firms, the strategy of subsidizing through complementarities is not an option. Consequently, SMEs are hindered by entry barriers due to the first mover's data network effects and a lack of resources (Spithoven, Vanhaverbeke, & Roijakkers, 2013). Hence, the research question addressed by this study is as follows: *What pathway can SMEs follow to enter a market with entry barriers that arise from incumbents' data network effects?*

3. Methods

Because there is no theory explaining how an SME can overcome the entry barriers raised by the activation of data network effects from the use of machine learning technology, epistemological discovery was sought (Locke, 2011) through an abductive research mode (Tavory & Timmermans, 2014). When existing theories cannot account for a given phenomenon (here, entry pathways), a deductive approach may limit inquiry by imposing predefined categories and their relations on the conception of the focal phenomenon. In contrast, a purely inductive mode of inquiry ignores previous knowledge to guide exploration (Behfar & Okhuysen, 2018). Abduction, regarded here as a combination of induction and deduction, aims to avoid the shortcomings of induction and deduction. It does so by allowing for openness to the discovery of new characteristics and relationships, while being able to link discoveries back to some pre-existing body of theory (Tavory & Timmermans, 2014). To this end, an exploratory longitudinal study of a start-up was performed for the purposes of theory building (Eisenhardt & Graebner, 2007). This approach is especially suitable for process questions such as the present research question (Eisenhardt, Graeber, & Sonenhein, 2016). The exploratory characterization of the start-up's path to the marketplace showed that the start-up had successfully provided a novel service that targeted stakeholders overlooked by competitors. Those stakeholders were examined in a dedicated follow-up study. The goal of that study was to confirm stakeholders' satisfaction with the novel services identified in the prior exploration of the entry process. A comprehensive description of that study and its results are detailed in the Appendix A.

3.1. The empirical setting

The setting was a start-up in the hemophilia healthcare sector in a Nordic country. The data spanned the period 2015 to 2021. This setting was chosen thanks to the opportunity that emerged during another empirical study (Climent & Haftor, 2021), where a professional relationship was developed with a senior manager at one of the studied firms. In 2017, the manager created this start-up, referred to here as "Alfa". Thus, the sampling in this study was both opportunity-driven (Patton, 1990) and theoretically motivated because the characteristics of the studied phenomenon matched the targeted theoretical gap (Eisenhardt & Graebner, 2007).

3.2. Data sources

Alfa's entry process was studied with multiple data sources. Multiple data collection methods were used to gather both qualitative and quantitative data (for the latter, see the Appendix A). Semi-structured interviews were used to interview representatives of the focal firm and other stakeholders. These stakeholders included patients and their relatives, a patient interest organization for hemophilia patients, regional health agencies, the national Dental and Pharmaceutical Products

Table 1
Specification of the study's interviews and interviewees (denoted "stakeholders") from mid-2017 to end of 2021.

Stakeholder	No. of representatives	No. of interviews	Length	
			Mean min.	SD min.
Employees at Alfa (focal firm)	13	164	56	17
Patients	23	41	41	7
Relatives	11	19	22	6
Patient interest organization	6	27	46	9
Hemophilia treatment center	14	57	63	22
Dental and Pharmaceutical Products Benefit Agency	5	21	44	19
Regional agencies	6	18	51	11
Pharmaceutical firm A	2	4	56	3
Pharmaceutical firm B	2	5	46	7
Pharmaceutical firm C	3	8	71	13

Beneficiary Agency, the staff at three hemophilia treatment centers, and representatives from three pharmaceutical firms that actively promote hemophilia products. Table 1 provides a summary of the interviewees and interviews. Follow-up emails were sent to gather various details with specific inquiries. Statistics on the sales of hemophilia pharmaceutical products were gathered, together with archival information regarding the market and its regulations and practices, as well as information about hemophilia diseases and treatments.

The interviews at the focal firm targeted all of the company's 13 employees. The informants' historical and current professional profile and role were covered, followed by detailed specifications of the firm's intentions and rationales, as well as the actual activities conducted in relation to those intentions. The interviews typically ended with further inquiry into the items that arose during the interview. Nearly all employees were interviewed multiple times, and several other stakeholders were interviewed. The focus was on both understanding the role of each stakeholder in the entry of the focal firm and triangulating the data from other sources, including the representatives of the focal firm.

Several measures ensured the validity and reliability of the data. The first was the use of multiple sources from both within and outside the focal firm at different levels of the organizational hierarchy. Second, the anonymity of the informants encouraged them to speak with candor. Third, questioning was non-directive. Fourth, event tracking was used to ask informants to give chronological details of key events that had happened and were planned in order to build scenarios. The pharmaceutical industry is highly regulated within the European Union, so each country's dedicated authorities gather detailed data on the real sales and prices of every pharmaceutical product. Such data is freely available upon request. Thus, all market-related data for hemophilia products in the focal Nordic County was obtained.

3.3. Data analysis and theory development

Research database software was used to store the data. One author started by constructing a comprehensive case scenario of the events covering the three phases of the focal firm: pre-establishment, establishment, and market entry. The other author critically reviewed this scenario. When details were missing or unclear, follow-up data gathering provided clarifications. This process was iterated several times. Simultaneously, the other author conducted atheoretical coding of the collected data, followed by grouping into themes. These themes were subsequently critically reviewed by the first author. This process was iterated several times to update the themes. The case scenario was then compared with the themes, which generated insights into the pattern of market entry of the focal firm and its potential underlying logic. Thereafter, this pattern was compared with the theory of data network effects, which involved multiple comparisons that eventually produced an updated pattern. This updated pattern is reflected in the emergent theory presented below (Eisenhardt et al., 2016; Klein & Myers, 1999). In particular, the various actors, along with their interests and actions, were carefully mapped. This mapping enabled identification of a key theoretical anomaly in the case. Detection of this anomaly initiated retroduction back to the assumed theoretical framework. This inquiry in turn led to a resolution and the formulation of the theoretical contribution presented here. Overall, "*successful market entry*" is used as the dependent variable (Bourreau, Sun, & Verboven, 2021; Park, 2020).

4. Results

To answer the present research question, the empirical case study is presented first in atheoretical mode and then in a theory-generation process that results in new theoretical insights.

4.1. The case

The focal firm's local market was healthcare services in the hemophilia products market, specifically a disease management system for hemophiliacs. Between 2005 and 2014, six international pharmaceutical firms introduced novel hemophilia products to the national marketplace. A senior manager at one of these firms spotted a novel business opportunity and left employment in mid-2017 to establish a new firm, Alfa. The aim was to offer a new kind of disease management system that included the use of machine learning technology. The subsequent events that led to Alfa's successful market entry are separated into three phases. The first is *pre-establishment*, between 2010 and 2017, where the incumbent pharmaceutical firms launched their hemophilia pharmaceuticals, followed by complementary digital services in the form of smartphone applications. The second phase is *establishment*, from mid-2017 until early 2019, when Alfa was created and developed its key offering. The third phase is successful *market entry* from mid-2019 until the end of 2021. During this phase, the firm received a breakthrough contract from regional health agencies. Before detailing these three phases, an introduction to hemophilia is given to provide background on the rationale for Alfa's market entry.

4.1.1. Hemophilia

Hemophilia is a rare bleeding disorder in people where a lack of clotting factor causes blood not to clot properly. This lack of clotting factor can cause internal and external bleeding, leading to damage or death if not stopped (Franchini, 2013). Crucially, bleeding may occur both spontaneously and as a result of injury. Internal bleeds are common and typically occur in joints and muscles, although sometimes in the brain. If not stopped rapidly, internal bleeds can cause death or produce swelling and joint destruction (hips, knees, ankles, and shoulders) generating chronic pain and requiring surgery. There are two types of hemophilia: Hemophilia A is caused by low levels of factor VIII, whereas Hemophilia B is caused by low levels of factor IX. Alfa's services focus on Hemophilia A patients because of the larger patient and data volume (Franchini, 2013).

Individuals with moderate and severe hemophilia must be treated with special treatment that infuses or injects replacement clotting factor into a vein. Lifelong treatment with replacement factor is necessary. If done properly, individuals can enjoy a high quality of life, although with an adapted lifestyle. The replacement factor can be provided to hemophiliacs on demand or through prophylaxis. Prophylaxis means that the replacement factor is injected regularly, typically two or three times a week, depending on a number of factors. On-demand treatment is conducted to stop ongoing bleeding. In most developed countries, a prophylaxis-based treatment is pursued for severe and moderate hemophiliacs, accompanied with on-demand treatment when needed to stop ongoing bleeds (Shrestha, Eldar-Lissai, Hou, Lakdawalla, & Batt,

Table 2
Introduction of pharmaceutical products and related digital disease management systems by hemophilia health providers and the focal firm Alfa.

Introduction	Year	Patient	Relatives	PIO	HTC	RA	BA	MPA
Digital services: ML-enabled	2021	1, 3, 4, SME	SME	SME	SME	SME	SME	SME
	2020	2, 5				SME		
	2019	6, SME			SME			
Digital services: Not ML-enabled	2018		1, 4	1, 4		5	5	5
	2017	1, 4	2, 5	3, 6	1, 4	2	2	2
	2016	3, 6	2, 5	2, 5	3, 6			
	2015	2, 5			2, 5			
Pharmaceutical products	2014–2005	2005: 1; 2009: 2, 6; 2010: 4; 2012: 3; 2013: 5.						

Notes: PIO = patient interest organization; HTC = hemophilia treatment center; RA = regional agencies; BA = Dental and Pharmaceutical Products Beneficiary Agency; MPA = Medical Products Agency; 1, 2, 3, 4, 5, 6 = six different product brands; SME = the case start-up "Alfa"; ML-enabled = machine learning technology enabled digital services.

2017). This approach not only reduces the risk of fatality and major injuries but also substantially improves quality of life for hemophiliacs and reduces healthcare costs associated with damages caused by bleeds. The challenges of prophylaxis-based treatment include the need to determine exactly how much replacement factor to provide and when, which depends on a number of aspects. Too little replacement factor still exposes hemophiliacs to great danger, whereas too much drives the treatment costs up. These costs are already high, often on average slightly more than 100,000 Euro per patient per year (Shrestha et al., 2017).

Hemophiliacs are treated at special hemophilia treatment centers (HTCs). These HTCs offer diagnosis and customized lifelong treatment regimes. They have several specialized experts working with each patient: hematologists, dentists, nutritionists, physiotherapists, psychologists, and surgeons. Hemophiliacs are advised to follow a lifestyle that reduces risk and promotes their health in terms of type of work and leisure activities, travel, and dietary practices, as well as their intake of other medication (Lamiani, Bigi, Mancuso, Coppola, & Vegni, 2017).

4.1.2. *Pre-establishment*

Table 2 shows the temporality of the introduction of replacement factors and related digital disease management systems by the studied firms. The six pharmaceutical product brands were introduced between 2005 and 2013. Digital disease management systems that did not use machine learning techniques were introduced from 2015 to 2018 aimed at various stakeholders. Machine-learning-based services were introduced between 2019 and 2021. The focal start-up (Alfa) introduced its disease management system, which uses machine learning techniques, in 2020, expanding it in 2021.

Between 2005 and 2013, six international firms each introduced a new generation of pharmaceutical products that offered replacement factor for hemophiliacs with Hemophilia A (Table 2). These products are highly similar in terms of their efficacy and safety profiles (Climent & Haftor, 2021). The introduction of such prescription drugs to the market is highly regulated. It starts with medical approval of a given drug by the Medical Products Agency. This approval is followed by product price setting by the Dental and Pharmaceutical Products Benefits Agency because public medical insurance covers most of the costs associated with a citizen's medical healthcare. This price setting by the Dental and Pharmaceutical Products Benefits Agency intends to eliminate price-based competition between pharmaceutical firms, meaning that products that target the same medical indication and that have similar efficacy profiles are assigned similar prices. Direct-to-consumer marketing for prescription-based pharmaceuticals is prohibited in the European Union. Therefore, pharmaceutical firms' key channel to promote their products are sales representatives, who regularly meet with the physicians that prescribe their drugs to detail the product's profile. Prescribers are key gatekeepers because they choose which products to prescribe to patients. However, drug selection is influenced by an official list of recommended drugs for each medical indication. This list is issued by each regional agency. The Nordic country where Alfa is located is geographically and administratively organized into 21 regions. Each region is managed by a regional agency that is responsible for healthcare financing and provision to its citizens. The drug selection process conducted by prescribers may also be influenced by lobbying from a patient interest organization and specific patients and relatives.

In 2015, six pharmaceutical firms shared the market. Two dominant firms had between 27 % and 30 % of the market share, two intermediate firms had 12 % to 15 %, and two lagging firms had 7 % to 9 % (Climent & Haftor, 2021). Because the six firms promoted six similar hemophilia products and price-based and marketing-driven competition was not possible, they launched complementary digital disease management systems in the form of dedicated software applications (apps) for smartphones. At the time of the study, the penetration of the use of smartphones was nearing 100 % of the country's population, and hemophilia patients already kept dairies. In these diaries, patients record their medical history (experience of bleeds: when, where, and why), the use of replacement factor and other medicines, and their lifestyle and nutrition habits, as well as their health status and well-being. These diaries are reported regularly to physicians, who analyze their content to adjust the replacement factor use and other kinds of treatment, as well as lifestyle recommendations.

In 2015, two of the six firms launched their app, with digital services targeting patients and HTC staff. Two other firms followed with a similar launch in 2016. The two pioneers then expanded their digital services toward other stakeholders (relatives and the patient interest organization). The two remaining firms launched their first app in 2017, whereas the other four firms, which had already launched their apps, expanded their digital services toward other stakeholders. Eventually, in 2018, the six firms were offering similar digital services via their apps to the same stakeholders (Table 2).

The content of the digital services provided by these firms can be conceived in terms of the activity chain of a hemophilia patient's disease management. Besides the functionality for (i) user profile registration, they offer functionalities for (ii) diagnosis, prognosis, and treatment profile registration by the HTC, (iii) patient reminders for replacement factor injections at the right time and dosage, as well as reminders for other activities including visits to the HTC, (iv) patients' registration of actual bleeds (where, when, and how much), (v) the actual treatment conducted, such as replacement factor injections (when, where, dosage, and treatment mode in the form of prophylaxis, on-demand, other medications, and measures), (vi) data records of lifestyle, (vii) lifestyle support (dietary recommendations, addresses of pharmacies and HTC centers when traveling, and recommended physical exercise sites), and (viii) customized report generation for HTC staff, patients, and relatives (Climent & Haftor, 2021).

4.1.3. *Establishment*

In 2016, one of the six pharmaceutical firms started to develop a recommendation functionality for replacement factor intakes for each individual who used their product. That recommendation functionality was aimed to be provided through the patient app. Patients using this firm's replacement factor registered numerous data points daily. This data was analyzed with machine learning techniques to identify patterns and then derive predictions to generate personalized recommendations for replacement factor intakes in terms of when and how much to inject. The timing and amount of appropriate replacement factor intake is co-conditioned by a number of factors (Aoki, Sakamoto, & Furutani, 2015; Lopes, Rios, Nogueira, & Mello, 2021). These factors include the individual's level of natural coagulation factor and volume and timing of previous intakes, physiological characteristics such as age, body mass, and metabolism, and behavioral factors such as risk profile, dietary and physical exercise habits, and the use of other medications (Franchini, 2013). The aim of the pharmaceutical firm that pioneered this recommendation functionality was to provide patients with more precise recommendations for replacement factor intakes than those that patients made on their own.

For various reasons (technological, medical, regulatory, and administrative) the development of this functionality was heavily delayed, and the pharmaceutical firm did not introduce it until the second half of 2019 (Table 2). One of the senior managers at that firm recognized a market opportunity with the introduction of app-mediated digital disease management system for hemophiliacs. The app should include a function for recommended replacement factor intake. This recommendation would be based on large data sets analyzed with machine learning techniques. Moreover, it would be unlimited to a specific replacement product brand. The idea was to target all hemophiliacs with one disease management system rather than having each firm target the patients who used its specific replacement factor products. As reported by the CEO of Alfa in one of the interviews:

> *Targeting all patients with one app only in a marketplace potentially allows for a larger user base than each pharmaceutical firm, which targets only its replacement factor users. This larger user base can provide much larger data sets that are important for the quality of predictions and therefore the accuracy of recommendations provided to users.*

This CEO had nearly 20 years of experience in marketing and sales in pharmaceutical firms across Nordic countries. In the first half of 2017, the CEO resigned to establish a venture, Alfa, to launch a digital disease management system to hemophiliacs that was independent of any product brand. The CEO secured financial capital from two private investors and employed a family member (daughter), who graduated from engineering school with a master's degree in computational sciences, to develop the machine learning technology. Several software engineers and usability designers were also recruited, together with a senior key account manager and a senior communication professional, both from the CEO's former pharmaceutical firm. The number of employees grew to 13 in 2021. The initial priority was to develop a prototype and test it, or, as the CEO of Alfa explained:

> *I believed, and still believe, that a key to success with this business is to develop a superior app, both in terms of its functionality and its usability. And as the existing apps are not that well designed, we have a good chance to succeed with ours.*

Within a year, Alfa had developed a full initial version of its app for hemophiliacs. Patient testing started in the second half of 2018, and it was introduced to the market at the beginning of 2019. This market launch was not successful because the initial uptake among patients was very low and stagnated quickly. This outcome came as a surprise, as noted by Alfa's Market Director and head of the market launch:

Well, we clearly underestimated how challenging it is to launch an app of this kind. We only had experience from launching pharmaceutical products.

Alfa conducted a number of in-depth interviews with patients to learn more about their needs with regard to the use of the app. These interviews revealed two kinds of limitations that their digital service would benefit from. One was the user design of the app. The other was the accuracy of recommendations for patients' intake of replacement factor. The user design was relatively uncomplicated for Alfa to improve. However, the improvement of prediction accuracy represented a major challenge because of the lack of data to generate more accurate recommendations. At that time, each pharmaceutical company could access data from patients that used its products, whereas each HTC could access all the data on its patients. There were three HTCs in the country, implying that no single actor could access the data for all patients. Alfa's management (CTO) understood this issue:

To improve the accuracy of our predictions, we needed to access all data available at the three hemophilia treatment centers, but we lacked the power to do so.

4.1.4. Market entry

The challenge of data access made Alfa radically modify its strategic intentions, from initially targeting patients to targeting regional agencies. The CEO described this shift:

As we realized that regulatory agencies had the power to provide us with all available patient data, we also knew that they hold a keen interest in the actual uses of pharmaceuticals and their costs.

Within three months, Alfa had established close dialogues with the three regional authorities where the country's three HTCs are located. These three authorities had the power to decide on the use of patient data. Alfa's message to them was that if Alfa was provided with data on all hemophiliacs' use of replacement factor, it had the knowledge and technology to analyze the actual use of the replacement factor continuously so as to detect potential suboptimal use of replacement factor. The acquisition and use of knowledge is crucial in the process of enterprise development (Zhao, Wen, Zhou, Liu, Yu & Xu, 2022). This analysis would provide the regional agencies with specific information on unnecessary costs due to overdosages of replacement factor and possible underuse of replacement factor that may cause serious health issues and additional healthcare costs. The more data that was provided and updated, the better the analysis would be at pinpointing potential areas to reduce overuse and underuse of replacement factor. Alfa was successful, thanks to its know-how, as reported by the Section Manager of the regional agency (top) and Alfa's Market Director (bottom):

> *When they presented the idea to us, it did not take us long to decide to say yes to a test as they clearly had detailed know-how and technology to do the kind of analysis that we are supposed to do but lack capability to do it. We were rather surprised how quickly the three regional agencies agreed to provide us with all the data to test our service. Initially, we were a bit worried that if we presented the idea to them, they would do the analysis themselves. From our discussions with them, we realized that they do not have that technical capability and that they need us.*

After initial tests, Alfa received two-year contracts in 2020 with all but one regional authority in the country. Toward the end of 2021, Alfa succeeded in receiving a similar contract from the national Dental and Pharmaceutical Products Beneficiary Agency. Their interest was to understand the actual dynamics of this type of pharmaceutical use as a foundation for forthcoming price renegotiations with the six pharmaceutical firms that provide the replacement factor. Such renegotiations are common practice every three to five years.

Alfa then updated its recommendation functionality for patients' usage of replacement factor and relaunched its updated patient app in 2021. This update also offered some specific functionalities for patients' relatives, such as parents. This time, its app-mediated services were also provided to and promoted by the hemophiliacs' patient interest organization, which enabled a faster user uptake than with the first launch attempt. Alfa now mainly targeted patients who were new to the use of replacement factor or were switching from one brand to another. Because there were patients who used the same product brand for many years, it would take considerable time before Alfa could gain a large market share of digital services for hemophiliacs.

On the revenue side, Alfa's initial plan was to charge a small subscription fee to patients who used their services. That plan implied that

Alfa would need to expand internationally before it could earn enough revenues for a positive financial result. Alfa's broadening of target groups to include regional authorities and the Dental and Pharmaceutical Products Beneficiary Agency, however, provided large enough and stable revenue sources so that Alfa did not need to charge a user subscription fee, which could deter some patients from using the service. Alfa offered its disease management system services to hemophiliacs for free. At the end of this study, in 2021, two of the pharmaceutical firms that were lagging in their digital services provision to patients, announced that they would stop providing such services and would recommend that their patients use Alfa's services, which added a significant volume of users to Alfa's services.

4.2. Theory generation

This section outlines some lessons from the detailed Alfa case and pursues an abductive process of theoretical development. The starting point is theorization on conventional product market strategies and more recent complementarities before the focus turns to the theory of data network effects.

First, the six pharmaceutical firms had difficulties with differentiation in terms of their product market strategies. Here, these strategies refer to a firm's actions to achieve and maintain a competitive advantage through positioning in product markets (Porter, 1985). The efficacy profiles of the available replacement factors were similar. This similarity led firms, one by one, to attempt to differentiate themselves by introducing digital services in the form of a patient-centric app that provided various disease management functionalities. This measure focused on the realization of positive complementarities between a firm's offerings. The idea is that the value of A is greater in the presence of B than other alternatives such as A by itself or A and C together (Brandenburger & Nalebuff, 1996; Ennen & Richter, 2010). This attempt at differentiation through complementarities was unsuccessful because the pioneering firms' introduction of digital services was quickly imitated by followers, which restored the status quo of a lack of differentiation. By introducing the digital disease management systems, the six pharmaceutical firms obtained data network effects to a minor extent only because of their small user base given that their users were restricted to patients who used their products. Small businesses find it difficult to develop scale-up and scale-up advantages and, as a result, face barriers to innovation capabilities (Hilmersson, & Hilmersson, 2021).

Alfa's intention was to exploit a latent unaddressed market opportunity by activating stronger data network effects from reaching as many hemophiliacs as possible and not being limited to users of a specific brand of replacement factor, thereby acquiring larger data sets than any competing service. This approach provided patients with high-quality recommendations, thereby achieving lock-in of users of their services. However, the initial attempt failed for two known reasons. First, there were limitations in the user-centric design of Alfa's services. Second, Alfa was unable to provide accurate enough recommendations for replacement factor usage. This second reason for failure was due to the lack of data volume because of slow user adoption of the disease management service. In this case, Alfa did not succeed in activating strong data network effects (Gregory et al., 2021). This failure is predicted by the theory of data network effects, which holds that users' perceived value of a service is conditioned by the accuracy of the predictions used to make recommendations. This relationship is, in turn, co-conditioned by the required volume of data and by the user-centric design of such services.

To overcome this failure, Alfa targeted another potential user of the same data, namely the regional authorities and, subsequently, the national Dental and Pharmaceutical Products Beneficiary Agency. They were offered value from alternative analysis of the same data. By identifying overdosage and underdosage of the replacement factor, Alfa pinpointed sources of substantial unnecessary costs in terms of both wasted medication and extra healthcare provisions due to undermedication. The value for the regional authorities lay in being precisely informed about where their subsequent measures should be focused to reduce overmedication and undermedication. The value for the Dental and Pharmaceutical Products Beneficiary Agency came from their understanding of the actual usage of these products as a foundation for forthcoming price renegotiations with pharmaceutical firms. The key to this retargeting of stakeholders lay in the authority of these new users of the data analysis to grant access to all the available user data generated by hemophiliacs from their use of the various disease management systems. They had the authority to legitimize data access, and they did so because it gave them valuable insights. Once this access had been secured, Alfa enhanced the user-centricity of its services, improved the quality of replacement factor usage recommendations based on the larger set of data, and relaunched the disease management app to the patients more successfully.

This detailed exploration of Alfa's successful path of entry to the marketplace shows that a key enabler for that success was the provision of services to a broader range of stakeholders than those previously reached with existing services. To confirm the validity of this success factor, a follow-up confirmatory study was conducted. Representatives of the targeted stakeholders were asked about their perceptions of these services. The follow-up study is detailed in the Appendix A. The results confirm that all stakeholder groups are highly satisfied with Alfa's services.

This case shows that users of machine-learning-based services should be regarded as *multiple current and potential actors*. These users have varying needs. Crucially, multiple actors can receive various machine learning services derived from the same data sets. This study also shows the importance of *temporal unfolding*, where current and potential users co-evolve together with the services being provided. There are various interdependencies between actors, data sets, and services. In the context of market entry by SMEs, this study's results suggest that SMEs can successfully enter a market dominated by pioneering incumbents by securing legitimate access to existing user data needed to derive services through machine learning.

5. Discussion

The research question addressed in this study is as follows: *What pathway can SMEs follow to enter a market with entry barriers that arise from incumbents' data network effects?* The following insights emerge in response to this question.

5.1. Theoretical contributions

This study offers at least five contributions to theory advancement. The first regards the notion of the user of machine-learning-based services. The theory of data network effects holds that a single human individual uses such services. While this situation may be the case, this study shows that the notion of a user should be broader. The empirical evidence is that users may include several social actors, both individuals and organizations, which use the same services or different services generated from the same data sets. This finding leads to the notion of the multi-actor user, as opposed to the single-actor user currently conceived by the theory. The present study shows that different actors that use the same or different services generated from the same data sets can have a range of profiles, with diverging interests and thus different legitimacy standards (Suchman, 1995). Hence, this expansion and diversification of the notion of the user implies that the factor of *data privacy use*, which co-conditions perceived usefulness, must be updated as well. More specifically, it needs to be diversified to account for each user actor, where data privacy use requirements may vary with each actor. Similarly, the other co-conditioning factors of perceived usefulness (i.e., performance expectancy, effort expectancy, data quality and quantity, speed, and accuracy) may have different requirements for each user actor. These requirements may diverge, as this case shows. The divergent nature of the multi-actor user introduces varying interests and incentives. Alignments, misalignments, and conflicts of interests may all arise. This scenario complicates the present version of the theory of data network effects, which currently only considers single users their own individual cognitive apparatus.

The second contribution comes from the observation that stakeholder theory (Freeman, 1984) can shape the understanding of the dynamics of data networks. Multiple stakeholders condition the activation of such networks. A firm's stakeholders are understood as "any group or individual who can affect or is affected by the achievement of the organization's objectives" (Freeman, 1984: 46). Examples of stakeholders include customers, their customers, suppliers, employees, owners, investors, NGOs, and regulators. Stakeholder theory predicts that a firm's positive relations with stakeholders contribute to firm performance, as opposed to a situation with negative stakeholder relations (Berrone, Surroca, & Tribo, 2007; Godfrey, 2005; Hillman & Keim, 2001). Whereas the mainstream strategic management literature focuses on the competition (Barney, 1991; Porter, 1985), the stakeholder approach emphasizes the very existence of a firm's stakeholders and the need to cooperate with them, or, in the words of Freeman (1984: 65), "stakeholder interests are joint." Accordingly, value co-creation with stakeholders is at the heart of stakeholder theory (Freeman, Harrison, & Wicks, 2007). To establish positive relations with stakeholders, the firm should take actions to make stakeholders perceive that it treats them fairly and rewards them (Hosmer & Kiewitz, 2005). The present empirical case study validates these predictions. As beneficiary agencies legitimize the access and use of patient data, these agencies are themselves rewarded, as is Alfa. However, some studies suggest that deliberate stakeholder management does not apply to SME success (Perrini, 2006). These studies are contradicted by the current findings, which imply that an SME can benefit from deliberate stakeholder management. This study shows that the activation of data network effects may be conditioned by actors other than the direct users of a machine-learning-enabled digital service. Moreover, these other actors may play a crucial role in the success of the service and the firm that provides it. Therefore, the findings of this study suggest that the implications of stakeholder theory can lead to further developments in the theory of data network effects (Clough & Wu, 2022; Gregory et al., 2021). This conclusion may be especially relevant for understanding how the use of machine learning technology can enable (a) value co-creation (Freeman, Dmytriyev, & Phillips, 2021), (b) business innovation, given that stakeholder management tends to correlate with firms' strong learning (Harting, Harmeling, & Venkataraman, 2006; Heugens, van den Bosch, & van Riel, 2002; Roome & Wijen, 2005), and (c) an organization's establishment of social legitimacy (Heugens, van den Bosch, & van Riel, 2002), which is crucial for the activation of data network effects (Gregory et al., 2021).

The third contribution of this study also relates to the theory of data network effects. The current theory suggests an atemporal notion of data network effects, with regard to who the user is and what user needs must be fulfilled. However, the present study shows a different situation. There may in fact be multiple current and potential users of services derived from the same data sets. Because there may be several competing providers of machine-learning-based services, variations may emerge with regard to which user actors are provided with what services, by which providers, and when. As this case shows, several providers (pharmaceutical firms) initially targeted a certain user actor (patients). Next, the start-up attempted to target the same user actor in an approach that failed. The start-up subsequently targeted another user actor (authorities). This co-evolution illustrates the use of the same data sets, for similar or varying services, by a range of providers, with the start-up gaining access to a large volume of data and generating superior services. Altogether, the notion of the multi-actor user and the temporal understanding of services and users lead to an updated notion of the theory of data network effects to account for these two related features.

The fourth contribution pertains to the literature on market entry by SMEs. The current literature on bypassing entry barriers is built on the industrial organization literature, with three main approaches: higher differentiation, broader scope, or niche markets (Acs, 2006; Arendt, 2008; Demsetz, 1982; Pehrsson, 2009). While these market entry strategies work for certain situations, they were conceived before the use of machine learning technology in service provision. Therefore, they are prior to the emergence of data network effects. However, they are invalidated by the data network effects that lock in the user-base and lock out competitors, which lack the relevant user data to generate high-quality services. The present study shows that an SME can successfully bypass the entry barriers erected by incumbents' activation of data network effects. An SME can enter a market by accessing the user data employed by incumbents. This access can be achieved by targeting another stakeholder's latent needs by offering machine learning-based services that require the same data as that used by incumbents. This stakeholder must have the authority and willingness to grant and legitimize this data access (Dwivedi, Kapoor, & Chen, 2015; Zimmerman & Zeitz, 2002). This willingness can be incentivized by the SME's offering. Once this stakeholder's needs have been met by the new services that activate data network effects, the SME can focus on other stakeholders targeted by incumbents with services based on the same data. This study thus reveals a pathway for SMEs to enter a market with entry barriers based on incumbents' activation of data network effects.

The fifth contribution is a result of the interpretation of the case explored in this study. With multiple actors in the form of firms, authorities, professionals, and individuals, data network effects account for, explain, and predict the course of actions that unfolded. This finding is important because data network theory is a recent theoretical contribution with limited empirical support (Clough & Wu, 2022; Gregory et al., 2021). Examples include the prediction that low user centricity and low data volume will hinder service providers' ability to generate high perceived user value from machine-learning-based services. The investigated start-up failed with these features of user centricity and data volume, which contributed to its initial lack of success. Once the start-up managed to resolve these issues, it relaunched its services successfully. This study thus offers a unique independent validation of the proposed theory of data network effects.

5.2. Implications for practice

The results offer several insights for practitioners. First, managers of both SMEs and incumbents that intend to invest in machine-learning-based services should first consider whether they wish such investment to be strategic. To play a strategic role for firms, machine-learning-based services must have the potential to achieve data network effects in such a manner that users become demotivated to migrate to alternative providers, while competitors struggle to establish high-quality services because of a lack of data access. If these conditions cannot be met, then the planned initiative will probably lack the potential to play a strategic role.

Second, the question of who will use the potential services or who may become a user of existing services or a given dataset is important. In non-trivial situations, several actors may become users of services. Some will generate usage data to help with the provision of services, while others will not. In such multi-actor user situations, it is crucial to decide which users will be provided with services and in which order. There are two key kinds of stakeholders. Those who generate usage data for machine learning to produce the service and those who legitimize access to such usage data. For strategic investments in machine learning use, managers may therefore benefit from adopting a temporal perspective to conceive which actors should be targeted by which services, who will generate the crucial data, and how such data can be legitimately accessed. This planning may involve competitors that also seek to provide similar services to the same users, the potential choice of users, and data access. Such planning can help make the right decision at the right time.

Third, firms that do not pioneer a given machine-learning-based service may struggle to enter a market because the pioneer has access to a large, unique set of user data that can help produce a superior service. While larger firms may compete head-to-head and use complementary services to attract a user base, SMEs typically lack such resources and struggle to enter the market, even when they have superior analysis technology. The pathway to bypass such barriers is to identify one or more stakeholders that can grant legitimate access to user data. For example, they can establish whether actors with the powers to provide access to data may benefit from services that are generated from the same data set that is targeted by the SME. There may then be an incentive for these actors to grant access to such data.

Fourth, policymakers should reflect on the possibility of establishing anti-trust policies that grant open and secure access to all user data equally for all actors. Such a policy could eliminate entry barriers based on access to user data, thereby leading to increased innovation and competition, which is a key source of user value creation over time.

5.3. Limitations and future research

The results of this study are original and contribute to knowledge on this topic. However, the results are derived from the exploration of a single case. Moreover, this case is rooted in the pharmaceutical industry, which has specific characteristics and heavy regulations. Also, the study represents a situation in a Northern European country, which has stable, advanced institutions. These factors may condition the results, so this study should be replicated in both similar and different contexts. Alternative contexts should cover other industries and other countries with specific characteristics to compare the results with those presented here. Another limitation is that the study offers only one possible pathway for an SME to enter a market with entry barriers based on incumbents' data network effects. Future research should investigate whether there are other pathways that also enable market entry. Finally, the activation of data network effects seems to be co-conditioned by stakeholders. Future research should combine the relevant components of stakeholder theory with the theory of data network effects to characterize how stakeholders condition the success or failure of data network effect activation.

Table 3
List of questions in the survey of stakeholders.

1. Using the system would allow me to perform the tasks for which I use the system more quickly.
2. The use of the system would facilitate the tasks for which I use the system.
3. The system would be useful for me to achieve the goal of my tasks.
4. Achieving my goal would be difficult without the system.
5. The system allows me to perform the tasks for which I use the system quicker than before.
6. The use of the system helps me achieve my goal to a greater extent than previous systems.
7. The use of the system responds to my needs more efficiently than before.
8. The system supports critical aspects related to the purpose for which I use the system.
9. I will increase my efficiency in the tasks for which I use the system.
10. I will be more efficient in terms of resources (money, time, etc.) linked to achieving my goal in using the system.
11. I will spend less time doing routine tasks which I use the system.
12. The people close to me (colleagues, family, friends, partners, etc.) perceive me as competent.
13. Learning how to operate a system like this would not take much effort for me.
14. It would be easy for me to get a system like this to do what I want it to do.
15. My interaction with such a system would be clear and understandable.
16. It would be easy for me to become proficient in the use of such a system.
17. The use of the system takes too much time away from my normal tasks.
18. Working with the system is so cumbersome that it is difficult to understand what is going on.
19. The use of the system involves too much time for mechanical operations (e.g., data entry).
20. It takes too long to learn how to use the system to be worth the effort.
21. My interaction with the system is clear and understandable.
22. I think it is easy to get the system to do what I want it to do.
23. Overall, I think the system is easy to use.
24. Learning to operate the system is easy for me.

6. Conclusions

The recent adoption of machine learning technologies enables organizations to activate data network effects. These novel uses of digital technology raise a new kind of barrier to followers who wish to provide similar services. Such barriers are particularly troublesome to SMEs, which typically lack the resources to compete head-to-head with large incumbents. The question addressed by this research is therefore as follows: *What pathway can SMEs follow to enter a market with entry barriers that arise from incumbents' data network effects?* The results suggest that a viable pathway to market entry for an SME, in such situations, is to incentivize stakeholders of such services to grant access for the start-up to use the necessary user data. Hence, it is the ability to establish legitimacy of data access, not ownership, that is the key to successful market entry when providing machine-learning-based services. This finding adds to the literature on how SMEs can bypass entry barriers. By extension, this study also adds to the theory of data network effects in two ways, namely by upgrading the notion of a service user from a single individual to a multi-actor user and by explaining data network effects temporally, where actors' decisions and actions co-evolve to create and capture value.

CRediT authorship contribution statement

Darek M. Haftor: Writing – original draft, Supervision, Methodology, Conceptualization. **Ricardo Costa-Climent:** Writing – original draft, Visualization, Software, Formal analysis. **Samuel Ribeiro Navarrete:** Formal analysis Methodology, Visualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Survey design, method, and results

Exploration of the investigated start-up's successful market entry shows that a key success factor is the introduction of new services to a broader range of stakeholders. The introduced services are superior to existing services and competitors' services. This factor is crucial for the success of the start-up's successful entry path. The validity of this finding was investigated by an independent follow-up study. The results confirm the findings provided by the exploratory study. The survey design, methods, and results are now detailed.

Objective: The objective was to evaluate the perceptions of different users of Alfa's service. The service was a digital application (app).

Sample: Users were all the stakeholders of the app.

The start-up's employees and the different types of users of the app included patients, their relatives, the hemophiliac patient interest organization, medical professionals at the hemophilia treatment centers, regional agencies, the Dental and Pharmaceutical Products Beneficiary Agency, the national Medical Products Agency, and representatives from competing firms. A total of 50 surveys were completed.

Timeframe: Interviews were conducted in spring 2022. This timeframe was chosen because it allowed assessment of users' perception of the use of the app after a minimum amount of time had elapsed since its launch.

Basic constructs: The aim of the analysis was to identify stakeholders' satisfaction with the service provided by the start-up. The service was enabled by machine learning technology use. Therefore, key concepts from the theory of data network effects were adopted. These constructs

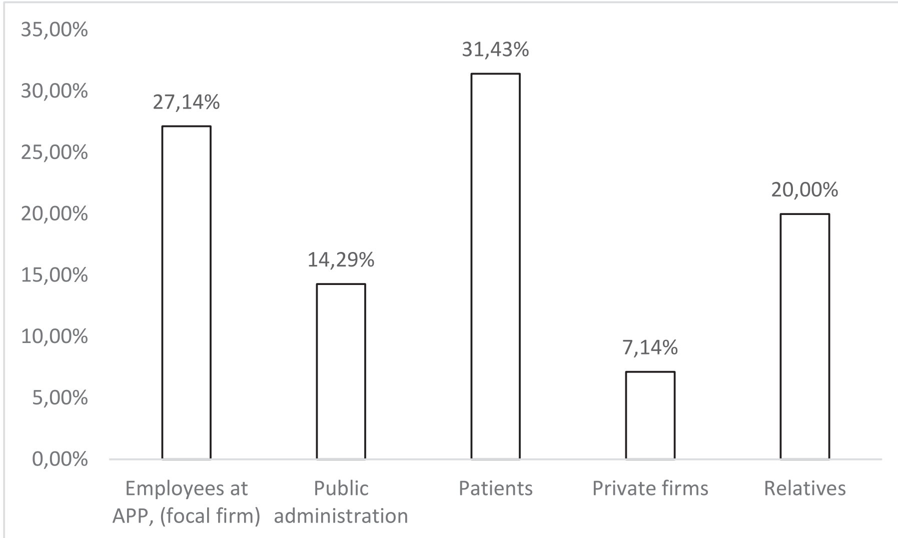

Fig. 1. Distribution by stakeholder group.

are performance expectancy and effort expectancy when using the app. Definitions of the concepts are provided later.

- Items related to "**performance expectancy**":
  - o *Perceived usefulness* (Davis, 1989).
  - o *Relative advantage* (Moore & Benbasat, 1991).
  - o *Expected outcome:* (Compeau & Higgins, 1995).
- Items related to "**effort expectancy**":
  - o *Perceived ease of use* (Davis, 1989; Davis et al., 1898).
  - o *Complexity* (Thompson et al., 1991).
  - o *Ease of use* (Moore & Benbasat, 1991).

Tool design: A survey was designed based on others validated in previous works (Venkatesh et al., 2003; Codrin & Mityko, 2012) and adapted to the current research aims. The instrument had two sections. The first collected identification data on participants (gender, age, and type of user). The second collected data on the remaining variables from 20 Likert-type items on 5-point scale (1–5) measuring performance expectancy and effort expectancy. In the survey, it was specified that the term "system" equates to a technology application that incorporates machine learning technology to learn from the data it acquires from users to make predictions and improvements to service offerings.

The survey was addressed to the stakeholders of the app (i.e., patients, employees, medical professionals, public administration, and so on). Therefore, the questions were expressed in terms of "…the tasks for which I use the app." The survey questions are listed in Table 3.

Material and statistical methodology: The statistical software used to conduct the statistical analysis was IBM SPSS Statistics 26.0.

Data analysis: Descriptive statistical analysis was used to analyze the data from respondents. For categorical variables, frequencies and percentages are shown. For continuous variables, means, standard deviations, medians, and interquartile ranges are shown. Comparisons between groups were made using a general linear model correcting for age. Robust covariances were used to handle the violation of assumptions. To test the reliability of the scales, Cronbach's alpha was used to measure internal consistency. All multiple comparisons were adjusted by Bonferroni correction. All tests are accompanied by the effect size estimator to enable interpretation of the results. For categorical variables, Cramer's V was used. For continuous variables Cohen's d was used. The criterion for classification of effect size is as follows: Cohen's d small ($d = 0.2$–0.4), medium ($d = 0.5$–0.8), and large (d = greater than 0.8); Cramer's V negligible (0.00–0.09), low (0.10–0.29), medium (0.30–0.49), and high ($\geq$0.50). The significance level used in the analyses was 5 % ($\alpha = 0.05$). Assuming no differences between groups, the *p* value is the probability that the results are due to chance. Any *p* value $<$ 0.05 indicates a statistically significant relationship. Conversely, a *p* value greater than or equal to 0.05 indicates no relationship.

Sample description: There were five types of respondents: one for each type of stakeholder (i.e., types of user of the app). Details are shown in Fig. 1.

There are significant differences between user types according to age: K-W (39.082, 4), *p* value = 0.000. Bonferroni-corrected post-hoc tests indicate that patients are significantly younger than the other users: $p < 0.05$, Cohen's $d = 0.89$–1.96. The start-up's (app's) employees are younger than relatives: $p < 0.05$, $d = 1.065$. Subsequent results are therefore corrected for age. See Table 4 for age distribution and Table 5 for gender distribution. There are no significant differences in gender by type of user (Table 5) (see Table 6).

Reliability and validity of scales: Cronbach's alpha, used to measure internal consistency, verifies the reliability of each scale construct by identifying whether the scale's reflective indicators are homogeneous. Therefore, it detects whether they are measuring the same latent variable. The optimum value is considered to be greater than or equal to 0.7.

The six factors met the criterion for Cronbach's alpha. Thus, the

Table 4
AGE according to USER TYPE (stakeholder type).

		Type of user						*p* value K-W
		Total	Employees at app (focal firm)	Public administration	Patients	Private firms	Relatives	
Age	N valid	70	19	10	22	5	14	0.000
	Average	35.30	34.58	41.70	24.23	41.40	46.93	
	Standard deviation	11.59	8.47	8.64	7.26	12.26	5.51	
	Median	33.50	31.00	41.50	22.50	49.00	46.00	
	Percentile 25	27.00	28.00	37.00	19.00	28.00	43.00	
	Percentile 75	44.00	40.00	47.00	28.00	51.00	49.00	

Table 5
GENDER according to USER TYPE (stakeholder type).

		Type of user												p value Chi
		Total		Employees at app (focal firm)		Public administration		Patients		Private firms		Relatives		
		Count	Column N %	Count	Column N %	Count	Column N %	Count	Column N %	Count	Column N %	Count	Column N %	
Gender	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.705
	Female	41	58.6 %	11	57.9 %	7	70.0 %	11	50.0 %	4	80.0 %	8	57.1 %	
	Male	29	41.4 %	8	42.1 %	3	30.0 %	11	50.0 %	1	20.0 %	6	42.9 %	

Table 6
Cronbach's alpha factor.

Factors	Cronbach's alpha
Perceived usefulness	0.931
Relative advantage	0.924
Expected outcome	0.761
Perceived ease of use	0.938
Complexity	0.930
Ease of use	0.947

analysis was conducted at the item level and at the dimension level, calculated as the average of the items in the dimension.

Results: Degree of agreement for each dimension.

Perceived utility

All respondents agree (around 25 %) or strongly agree (around 75 %) with the four items of perceived usefulness. The overall perceived usefulness (mean of the four items) does not differ by type of user (p = 0.652), as shown in Table 7. The average degree of agreement (1–5) shows no significant differences by user ($p > 0.05$).

Relative advantage

All respondents agree (around 25 %–30 %) or strongly agree (around 70 %–75 %) with the four items of relative advantage. The overall relative advantage (mean of the four items) does not differ by user type ($p = 0.652$), as shown by Table 8. Using the quantitative approach by calculating the average degree of agreement (1–5) also reveals non-significant differences by user (p greater than 0.05).

Expected outcome

All respondents agree (around 20 %–25 %) or strongly agree (around 75 %–80 %) with items 9, 10, 11 of expected outcome, as shown in Table 9. Item 12 is split between neutral (21 %), agree (37 %) and strongly agree (43 %). Public administrations and private companies have the most neutral opinions on this item. Relaxing the significance level to 10 % implies that the overall expected outcome has higher agreement among employees than among the other groups: F(1,966.4), $p = 0.096$, Cramer's V = 0.382. A quantitative approach by calculating the mean degree of agreement (1–5) ratifies the previous result: public administration has lower agreement with item 12 than patients, relatives, and employees: F(6,763.4), $p < 0.000$, Cohen's d = 0.980, 1.148, and 1.736, respectively. Hence, the overall expected outcome is higher in workers than in public administration: F(2.530,4), $p = 0.049$, Cohen's d = 1.090 (see Table 10).

Perceived ease of use

All respondents agree (around 25 %–35 %) or strongly agree (around 65 %–75 %) with the perceived ease of use items, as shown in Table 11 and Table 11. Item 13 has the lowest level of agreement. It also differs by user, with PPP employees agreeing much more than patients: F (2.563,4), $p = 0.049$, Cramer's V = 0.408. Items 14 and 15 also differ by

Table 7
PERCEIVED UTILITY according to TYPE OF USER.

		Type of user												p value
		Total		Employees at app (focal firm)		Public administration		Patients		Private firms		Relatives		
		N	%	N	%	N	%	N	%	N	%	N	%	
1_Using the system would allow me to perform the tasks for which I use the system more quickly.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.227
	Agree	17	24.3 %	2	10.5 %	4	40.0 %	8	36.4 %	1	20.0 %	2	14.3 %	
	Totally agree	53	75.7 %	17	89.5 %	6	60.0 %	14	63.6 %	4	80.0 %	12	85.7 %	
2_The use of the system would facilitate the tasks for which I use the system.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.883
	Agree	17	24.3 %	3	15.8 %	3	30.0 %	6	27.3 %	1	20.0 %	4	28.6 %	
	Totally agree	53	75.7 %	16	84.2 %	7	70.0 %	16	72.7 %	4	80.0 %	10	71.4 %	
3_The system would be useful for me to achieve the goal of my tasks.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.871
	Agree	12	17.1 %	2	10.5 %	2	20.0 %	4	18.2 %	1	20.0 %	3	21.4 %	
	Totally agree	58	82.9 %	17	89.5 %	8	80.0 %	18	81.8 %	4	80.0 %	11	78.6 %	
4_Achieving the goal for which I use the system would be difficult without the system.	Total	69	100.0 %	19	100.0 %	10	100.0 %	21	100.0 %	5	100.0 %	14	100.0 %	0.608
	Agree	18	26.1 %	3	15.8 %	4	40.0 %	6	28.6 %	1	20.0 %	4	28.6 %	
	Totally agree	51	73.9 %	16	84.2 %	6	60.0 %	15	71.4 %	4	80.0 %	10	71.4 %	
PERCEIVED UTILITY	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.661
	Agree	19	27.1 %	3	15.8 %	4	40.0 %	7	31.8 %	1	20.0 %	4	28.6 %	
	Totally agree	51	72.9 %	16	84.2 %	6	60.0 %	15	68.2 %	4	80.0 %	10	71.4 %	

Table 8
RELATIVE ADVANTAGE according to TYPE OF USER.

		Type of user												p value
		Total		Employees at app (focal firm)		Public administration		Patients		Private firms		Relatives		
		N	%	N	%	N	%	N	%	N	%	N	%	
5_The use of the system allows me to perform the tasks for which I use the system quicker than before.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.608
	Agree	20	28.6 %	3	15.8 %	4	40.0 %	8	36.4 %	1	20.0 %	4	28.6 %	
	Totally agree	50	71.4 %	16	84.2 %	6	60.0 %	14	63.6 %	4	80.0 %	10	71.4 %	
6_The use of the system helps me achieve the goal for which I use it to a greater extent than previous systems.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.768
	Agree	18	25.7 %	3	15.8 %	3	30.0 %	8	36.4 %	1	20.0 %	3	21.4 %	
	Totally agree	52	74.3 %	16	84.2 %	7	70.0 %	14	63.6 %	4	80.0 %	11	78.6 %	
7_The use of the system responds to my needs more efficiently than before.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.839
	Agree	20	28.6 %	4	21.1 %	3	30.0 %	8	36.4 %	1	20.0 %	4	28.6 %	
	Totally agree	50	71.4 %	15	78.9 %	7	70.0 %	14	63.6 %	4	80.0 %	10	71.4 %	
8_The system supports critical aspects related to the purpose for which I use the system.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.603
	Agree	24	34.3 %	4	21.1 %	4	40.0 %	10	45.5 %	1	20.0 %	5	35.7 %	
	Totally agree	46	65.7 %	15	78.9 %	6	60.0 %	12	54.5 %	4	80.0 %	9	64.3 %	
RELATIVE ADVANTAGE	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.661
	Agree	19	27.1 %	3	15.8 %	4	40.0 %	7	31.8 %	1	20.0 %	4	28.6 %	
	Totally agree	51	72.9 %	16	84.2 %	6	60.0 %	15	68.2 %	4	80.0 %	10	71.4 %	

Table 9
EXPECTED OUTCOME according to TYPE OF USER.

		Type of user												p value
		Total		Employees at app (focal firm)		Public administration		Patients		Private firms		Relatives		
		N	%	N	%	N	%	N	%	N	%	N	%	
9_I will increase my efficiency in the tasks for which I use the system.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.573
	Agree	16	22.9 %	2	10.5 %	2	20.0 %	6	27.3 %	1	20.0 %	5	35.7 %	
	Totally agree	54	77.1 %	17	89.5 %	8	80.0 %	16	72.7 %	4	80.0 %	9	64.3 %	
10_I will be more efficient in terms of resources (money, time, etc.) linked to achieving the goals for which I use the system.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.686
	Agree	15	21.4 %	2	10.5 %	3	30.0 %	6	27.3 %	1	20.0 %	3	21.4 %	
	Totally agree	55	78.6 %	17	89.5 %	7	70.0 %	16	72.7 %	4	80.0 %	11	78.6 %	
11_I will spend less time doing routine tasks for which I use the system.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.411
	Agree	18	25.7 %	2	10.5 %	3	30.0 %	7	31.8 %	1	20.0 %	5	35.7 %	
	Totally agree	52	74.3 %	17	89.5 %	7	70.0 %	15	68.2 %	4	80.0 %	9	64.3 %	
12_The people close to me (colleagues, family, friends, partners, etc.) will perceive me as competent.	Total	68	100.0 %	19	100.0 %	10	100.0 %	21	100.0 %	5	100.0 %	13	100.0 %	0.000
	Neither agree nor disagree	14	20.6 %	0	0.0 %	6	60.0 %	3	14.3 %	3	60.0 %	2	15.4 %	
	Agree	25	36.8 %	5	26.3 %	4	40.0 %	9	42.9 %	0	0.0 %	7	53.8 %	
	Totally agree	29	42.6 %	14	73.7 %	0	0.0 %	9	42.9 %	2	40.0 %	4	30.8 %	0.096
EXPECTED OUTCOME	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	
	Agree	28	40.0 %	2	10.5 %	6	60.0 %	10	45.5 %	3	60.0 %	7	50.0 %	
	Totally agree	42	60.0 %	17	89.5 %	4	40.0 %	12	54.5 %	2	40.0 %	7	50.0 %	

user type. They are rated better among employees than patients: F (40,263.8), $p = 0.000$, Cramer's V = 0.286; F(61,337.8), $p = 0.000$, Cramer's V = 0.310. Thus, overall perceived ease of use has higher agreement among employees than among patients: F(2.265,4), $p = 0.037$, Cramer's V = 0.382. A quantitative approach by calculating the average degree of agreement (1–5) ratifies the previous result. In addition, employees, private firms, and public administration rate the perceived ease of use better than patients and relatives. Hence, the overall perceived ease of use is higher in workers than in patients and relatives: F(3,043.4), $p = 0.023$, Cohen's d = 0.897 and 0.791, respectively.

Complexity

All respondents strongly disagree (around 75 %) or disagree (around 25 %) with the four complexity items, as shown in Table 12 and Table 13. There are no significant differences by type of user. Overall complexity (average of the four items) does not differ by type of user. A

Table 10
PERCEIVED EASE OF USE according to TYPE OF USER.

		Type of user												p value
		Total		Employees at app (focal firm)		Public administration		Patients		Private firms		Relatives		
		N	%	N	%	N	%	N	%	N	%	N	%	
13_Learning how to operate a system like this would not take much effort for me.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.049
	Agree	24	34.3 %	2	10.5 %	2	20.0 %	12	54.5 %	1	20.0 %	7	50.0 %	
	Totally agree	46	65.7 %	17	89.5 %	8	80.0 %	10	45.5 %	4	80.0 %	7	50.0 %	
14_It would be easy for me to get a system like this to do what I want it to do.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.000
	Neither agree nor disagree	1	1.4 %	0	0.0 %	0	0.0 %	1	4.5 %	0	0.0 %	0	0.0 %	
	Agree	19	27.1 %	2	10.5 %	1	10.0 %	8	36.4 %	1	20.0 %	7	50.0 %	
	Totally agree	50	71.4 %	17	89.5 %	9	90.0 %	13	59.1 %	4	80.0 %	7	50.0 %	
15_My interaction with such a system would be clear and understandable.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.000
	Neither agree nor disagree	1	1.4 %	0	0.0 %	0	0.0 %	1	4.5 %	0	0.0 %	0	0.0 %	
	Agree	21	30.0 %	2	10.5 %	1	10.0 %	10	45.5 %	1	20.0 %	7	50.0 %	
	Totally agree	48	68.6 %	17	89.5 %	9	90.0 %	11	50.0 %	4	80.0 %	7	50.0 %	
16_It would be easy for me to become proficient in the use of such a system	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.284
	Agree	19	27.1 %	2	10.5 %	2	20.0 %	8	36.4 %	1	20.0 %	6	42.9 %	
	Totally agree	51	72.9 %	17	89.5 %	8	80.0 %	14	63.6 %	4	80.0 %	8	57.1 %	
PERCEIVED EASE OF USE	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.037
	Agree	23	32.9 %	2	10.5 %	2	20.0 %	11	50.0 %	1	20.0 %	7	50.0 %	
	Totally agree	47	67.1 %	17	89.5 %	8	80.0 %	11	50.0 %	4	80.0 %	7	50.0 %	

Table 11
PERCEIVED EASE OF USE according to TYPE OF USER.

		Type of user						p value
		Total	Employees at app (focal firm)	Public administration	Patients	Private firms	Relatives	
13_Learning how to operate a system like this would not take much effort for me.	Valid N	70	19	10	22	5	14	0.016
	Mean	4.66	4.89	4.80	4.45	4.80	4.50	
	Standard Deviation	0.48	0.32	0.42	0.51	0.45	0.52	
	Median	5.00	5.00	5.00	4.00	5.00	4.50	
	Percentile 25	4.00	5.00	5.00	4.00	5.00	4.00	
	Percentile 75	5.00	5.00	5.00	5.00	5.00	5.00	
14_It would be easy for me to get a system like this to do what I want it to do.	Valid N	70	19	10	22	5	14	0.040
	Mean	4.70	4.89	4.90	4.55	4.80	4.50	
	Standard Deviation	0.49	0.32	0.32	0.60	0.45	0.52	
	Median	5.00	5.00	5.00	5.00	5.00	4.50	
	Percentile 25	4.00	5.00	5.00	4.00	5.00	4.00	
	Percentile 75	5.00	5.00	5.00	5.00	5.00	5.00	
15_My interaction with such a system would be clear and understandable.	Valid N	70	19	10	22	5	14	0.022
	Mean	4.67	4.89	4.90	4.45	4.80	4.50	
	Standard Deviation	0.50	0.32	0.32	0.60	0.45	0.52	
	Median	5.00	5.00	5.00	4.50	5.00	4.50	
	Percentile 25	4.00	5.00	5.00	4.00	5.00	4.00	
	Percentile 75	5.00	5.00	5.00	5.00	5.00	5.00	
16_It would be easy for me to become proficient in the use of such a system.	Valid N	70	19	10	22	5	14	0.235
	Mean	4.73	4.89	4.80	4.64	4.80	4.57	
	Standard Deviation	0.45	0.32	0.42	0.49	0.45	0.51	
	Median	5.00	5.00	5.00	5.00	5.00	5.00	
	Percentile 25	4.00	5.00	5.00	4.00	5.00	4.00	
	Percentile 75	5.00	5.00	5.00	5.00	5.00	5.00	
PERCEIVED EASE OF USE	Valid N	70	19	10	22	5	14	0.023
	Mean	4.69	4.89	4.85	4.52	4.80	4.52	
	Standard Deviation	0.44	0.32	0.34	0.46	0.45	0.50	
	Median	5.00	5.00	5.00	4.63	5.00	4.63	
	Percentile 25	4.25	5.00	5.00	4.00	5.00	4.00	
	Percentile 75	5.00	5.00	5.00	5.00	5.00	5.00	

Table 12
COMPLEXITY according to TYPE OF USER.

		Type of user												p value
		Total		Employees at app (focal firm)		Public administration		Patients		Private firms		Relatives		
		N	%	N	%	N	%	N	%	N	%	N	%	
17_The use of the system takes too much time away from my normal tasks.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.417
	Strongly disagree	54	77.1 %	18	94.7 %	7	70.0 %	16	72.7 %	4	80.0 %	9	64.3 %	
	Disagree	16	22.9 %	1	5.3 %	3	30.0 %	6	27.3 %	1	20.0 %	5	35.7 %	
18_Working with the system is so cumbersome that it is difficult to understand what is going on.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.250
	Strongly disagree	54	77.1 %	18	94.7 %	8	80.0 %	15	68.2 %	4	80.0 %	9	64.3 %	
	Disagree	16	22.9 %	1	5.3 %	2	20.0 %	7	31.8 %	1	20.0 %	5	35.7 %	
19_The use of the system involves too much time for mechanical operations (e.g., data entry).	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.264
	Strongly disagree	51	72.9 %	18	94.7 %	7	70.0 %	13	59.1 %	4	80.0 %	9	64.3 %	
	Disagree	19	27.1 %	1	5.3 %	3	30.0 %	9	40.9 %	1	20.0 %	5	35.7 %	
20_It takes too long to learn how to use the system to be worth the effort.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.390
	Strongly disagree	53	75.7 %	18	94.7 %	7	70.0 %	15	68.2 %	4	80.0 %	9	64.3 %	
	Disagree	17	24.3 %	1	5.3 %	3	30.0 %	7	31.8 %	1	20.0 %	5	35.7 %	
COMPLEXITY	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.361
	Strongly disagree	57	81.4 %	18	94.7 %	8	80.0 %	18	81.8 %	4	80.0 %	9	64.3 %	
	Disagree	13	18.6 %	1	5.3 %	2	20.0 %	4	18.2 %	1	20.0 %	5	35.7 %	

Table 13
COMPLEXITY according to TYPE OF USER.

		Type of user						p value
		Total	Employees at app (focal firm)	Public administration	Patients	Private firms	Relatives	
17_The use of the system takes too much time away from my normal tasks.	Valid N	70	19	10	22	5	14	0.309
	Mean	1.23	1.05	1.30	1.27	1.20	1.36	
	Standard Deviation	0.42	0.23	0.48	0.46	0.45	0.50	
	Median	1.00	1.00	1.00	1.00	1.00	1.00	
	Percentile 25	1.00	1.00	1.00	1.00	1.00	1.00	
	Percentile 75	1.00	1.00	2.00	2.00	1.00	2.00	
18_Working with the system is so cumbersome that it is difficult to understand what is going on.	Valid N	70	19	10	22	5	14	0.201
	Mean	1.23	1.05	1.20	1.32	1.20	1.36	
	Standard Deviation	0.42	0.23	0.42	0.48	0.45	0.50	
	Median	1.00	1.00	1.00	1.00	1.00	1.00	
	Percentile 25	1.00	1.00	1.00	1.00	1.00	1.00	
	Percentile 75	1.00	1.00	1.00	2.00	1.00	2.00	
19_The use of the system involves too much time for mechanical operations (e.g., data entry).	Valid N	70	19	10	22	5	14	0.134
	Mean	1.27	1.05	1.30	1.41	1.20	1.36	
	Standard Deviation	0.45	0.23	0.48	0.50	0.45	0.50	
	Median	1.00	1.00	1.00	1.00	1.00	1.00	
	Percentile 25	1.00	1.00	1.00	1.00	1.00	1.00	
	Percentile 75	2.00	1.00	2.00	2.00	1.00	2.00	
20_It takes too long to learn how to use the system to be worth the effort.	Valid N	70	19	10	22	5	14	0.238
	Mean	1.24	1.05	1.30	1.32	1.20	1.36	
	Standard Deviation	0.43	0.23	0.48	0.48	0.45	0.50	
	Median	1.00	1.00	1.00	1.00	1.00	1.00	
	Percentile 25	1.00	1.00	1.00	1.00	1.00	1.00	
	Percentile 75	1.00	1.00	2.00	2.00	1.00	2.00	
COMPLEXITY	Valid N	70	19	10	22	5	14	0.023
	Mean	1.24	1.05	1.28	1.33	1.20	1.36	
	Standard Deviation	0.39	0.23	0.42	0.38	0.45	0.50	
	Median	1.00	1.00	1.00	1.25	1.00	1.00	
	Percentile 25	1.00	1.00	1.00	1.00	1.00	1.00	
	Percentile 75	1.50	1.00	1.50	1.50	1.00	2.00	

Table 14
EASE OF USE according to TYPE OF USER.

		Type of user												p value
		Total		Employees at app (focal firm)		Public administration		Patients		Private firms		Relatives		
		N	%	N	%	N	%	N	%	N	%	N	%	
21_My interaction with the system is clear and understandable.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.359
	Agree	17	24.3 %	1	5.3 %	3	30.0 %	7	31.8 %	1	20.0 %	5	35.7 %	
	Totally agree	53	75.7 %	18	94.7 %	7	70.0 %	15	68.2 %	4	80.0 %	9	64.3 %	
22_I think it is easy to get the system to do what I want it to do.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.333
	Agree	18	25.7 %	1	5.3 %	4	40.0 %	7	31.8 %	1	20.0 %	5	35.7 %	
	Totally agree	52	74.3 %	18	94.7 %	6	60.0 %	15	68.2 %	4	80.0 %	9	64.3 %	
23_Overall, I think the system is easy to use.	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.406
	Agree	14	20.0 %	1	5.3 %	3	30.0 %	4	18.2 %	1	20.0 %	5	35.7 %	
	Totally agree	56	80.0 %	18	94.7 %	7	70.0 %	18	81.8 %	4	80.0 %	9	64.3 %	
24_Learning to operate the system is easy for me.	Total	68	100.0 %	17	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.241
	Agree	17	25.0 %	1	5.9 %	3	30.0 %	6	27.3 %	3	60.0 %	4	28.6 %	
	Totally agree	51	75.0 %	16	94.1 %	7	70.0 %	16	72.7 %	2	40.0 %	10	71.4 %	
EASE OF USE	Total	70	100.0 %	19	100.0 %	10	100.0 %	22	100.0 %	5	100.0 %	14	100.0 %	0.360
	Agree	16	22.9 %	1	5.3 %	3	30.0 %	6	27.3 %	1	20.0 %	5	35.7 %	
	Totally agree	54	77.1 %	18	94.7 %	7	70.0 %	16	72.7 %	4	80.0 %	9	64.3 %	

Table 15
EASE OF USE according to TYPE OF USER.

		Type of user						p value
		Total	Employees at app (focal firm)	Public administration	Patients	Private firms	Relatives	
21_My interaction with the system is clear and understandable.	Valid N	70	19	10	22	5	14	0.316
	Mean	4.76	4.95	4.70	4.68	4.80	4.64	
	Standard Deviation	0.43	0.23	0.48	0.48	0.45	0.50	
	Median	5.00	5.00	5.00	5.00	5.00	5.00	
	Percentile 25	5.00	5.00	4.00	4.00	5.00	4.00	
	Percentile 75	5.00	5.00	5.00	5.00	5.00	5.00	
22_I think it is easy to get the system to do what I want it to do.	Valid N	70	19	10	22	5	14	0.238
	Mean	4.74	4.95	4.60	4.68	4.80	4.64	
	Standard Deviation	0.44	0.23	0.52	0.48	0.45	0.50	
	Median	5.00	5.00	5.00	5.00	5.00	5.00	
	Percentile 25	4.00	5.00	4.00	4.00	5.00	4.00	
	Percentile 75	5.00	5.00	5.00	5.00	5.00	5.00	
23_Overall, I think the system is easy to use.	Valid N	70	19	10	22	5	14	0.461
	Mean	4.80	4.95	4.70	4.82	4.80	4.64	
	Standard Deviation	0.40	0.23	0.48	0.39	0.45	0.50	
	Median	5.00	5.00	5.00	5.00	5.00	5.00	
	Percentile 25	5.00	5.00	4.00	5.00	5.00	4.00	
	Percentile 75	5.00	5.00	5.00	5.00	5.00	5.00	
24_Learning to operate the system is easy for me.	Valid N	68	17	10	22	5	14	0.161
	Mean	4.75	4.94	4.70	4.73	4.40	4.71	
	Standard Deviation	0.44	0.24	0.48	0.46	0.55	0.47	
	Median	5.00	5.00	5.00	5.00	4.00	5.00	
	Percentile 25	4.50	5.00	4.00	4.00	4.00	4.00	
	Percentile 75	5.00	5.00	5.00	5.00	5.00	5.00	
EASE OF USE	Valid N	70	19	10	22	5	14	0.046
	Mean	4.76	4.95	4.68	4.73	4.70	4.66	
	Standard Deviation	0.40	0.23	0.47	0.40	0.41	0.48	
	Median	5.00	5.00	5.00	5.00	4.75	5.00	
	Percentile 25	4.75	5.00	4.00	4.50	4.75	4.00	
	Percentile 75	5.00	5.00	5.00	5.00	5.00	5.00	

quantitative approach by calculating the average degree of agreement (1–5) suggests that employees perceive greater overall complexity than patients and relatives: F(3.309,4), $p = 0.023$, Cohen's d = 0.674 and 0.814, respectively.

Ease of use

All respondents agree (20 %–25 %) or strongly agree (75 %-80 %) with the four ease of use items, as shown in Table 14 and Table 15. There are no significant differences by type of user. Overall ease of use (mean of the four items) does not differ by type of user ($p = 0.360$). A quantitative approach by calculating the mean degree of agreement (1–5) shows that employees perceive greater overall ease of use than patients: F(2.577,4), $p = 0.046$, Cohen's d = 0.619.

Conclusions: Overall, the ratings of items across all dimensions are very positive. Most users agree or strongly agree with the questionnaire items. Opinions on *perceived usefulness*, *relative advantage*, and *complexity* do not differ significantly between user types. Opinions on *expected outcome* and *perceived ease of use* differ somewhat by type of user. In general, employees of the start-up rate these items better than patients and, in some cases, public administration and private employees. This analysis provides an exploration of the successful market entry path of a start-up that bypasses entry barriers due to incumbents' activated data network effects. This analysis confirms that the start-up has succeeded in providing a more valuable service offering to its stakeholders than the existing service offering.

References

Abbasi, K., Alam, A., Du, M. A., & Huynh, T. L. D. (2021). FinTech, SME efficiency and national culture: evidence from OECD countries. *Technological Forecasting and Social Change, 163*, Article 120454.

Adamides, E., & Karacapilidis, N. (2020). Information technology for supporting the development and maintenance of open innovation capabilities. *Journal of Innovation & Knowledge, 5*(1), 29–38. https://doi.org/10.1016/j.jik.2018.07.001

Afuah, A. (2013). Are network effects really all about size? The role of structure and conduct. *Strategic Management Journal, 34*(3), 257–273.

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines: The simple economics of artificial intelligence*. Harvard Business Press.

Alter, S. (2021). Understanding artificial intelligence in the context of usage: Contributions and smartness of algorithmic capabilities in work systems. *International Journal of Information Management*, Article 102392.

Amit, R., & Zott, C. (2001). Value creation in e-business. *Strategic Management Journal, 22* (6–7), 493–520.

Aoki, K., Sakamoto, M., & Furutani, H. (2015). Analysis of genetic disease haemophilia A by using machine learning. *Journal of Robotics, Networking and Artificial Life, 2*(2), 115–119.

Arendt, L. (2008). Barriers to ICT adoption in SMEs: How to bridge the digital divide. *Journal of Systems and Information Technology, 10*(2), 93–108.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120.

Behfar, K., & Okhuysen, G. A. (2018). Perspective—Discovery within validation logic: Deliberately surfacing, complementing, and substituting abductive reasoning in hypothetico-deductive inquiry. *Organization Science, 29*, 323–340.

Berrone, P., Surroca, J., & Tribó, J. A. (2007). Corporate ethical identity as a determinant of firm performance: A test of the mediating role of stakeholder satisfaction. *Journal of Business Ethics, 76*, 35–53.

Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quarterly*, 169–196.

Boudreau, K. J., & Jeppesen, L. B. (2015). Unpaid crowd complementors: The platform network effect mirage. *Strategic Management Journal, 36*(12), 1761–1777.

Bourreau, M., Sun, Y., & Verboven, F. (2021). Market entry, fighting brands, and tacit collusion: evidence from the french mobile telecommunications market. *American Economic Review, 111*(11), 3459–3499.

Boxu, Y., Xingguang, L., & Kou, K. (2022). Research on the influence of network embeddedness on innovation performance: Evidence from China's listed firms. *Journal of Innovation & Knowledge, 7*(3), Article 100210.

Brandenburger, A. M., & Nalebuff, B. J. (1996). *Co-opetition*. New York: Doubleday.

Brynjolfsson, E., & Saunders, A. (2009). *Wired for innovation: How information technology is reshaping the economy*. MIT Press.

Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge, MA: Harvard University Press.

Chae, H. C., Koh, C. E., & Prybutok, V. R. (2014). Information technology capability and firm performance: Contradictory findings and their possible causes. *MIS Quarterly, 38*(1), 305–326.

Climent, R. C., & Haftor, D. M. (2021). Business model theory-based prediction of digital technology use: An empirical assessment. *Technological Forecasting and Social Change, 173*, Article 121174.

Clough, D. R., & Wu, A. (2022). Artificial intelligence, data-driven learning, and the decentralized structure of platform ecosystems. *Academy of Management Review, 47* (1), 184–189.

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management, 60*, Article 102383.

Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly*, 189–211.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319–340.

Demsetz, H. (1982). Barriers to entry. *The American Economic Review, 72*(1), 47–57.

Doz, Y. L., & Hamel, G. (1998). *Alliance advantage: The art of creating value through partnering*. Harvard Business Press.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63–71.

Dutta, S., & Segev, A. (1999). Business transformation on the Internet. *European Management Journal, 17*(5), 466–476.

Dwivedi, Y. K., Kapoor, K. K., & Chen, H. (2015). Social media marketing and advertising. *The Marketing Review, 15*(3), 289–309.

Dwivedi, Y. K., Rana, N. P., Janssen, M., Lal, B., Williams, M. D., & Clement, M. (2017). An empirical validation of a unified model of electronic government adoption (UMEGA). *Government Information Quarterly, 34*(2), 211–230.

Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., & Williams, M. D. (2019). Reexamining the unified theory of acceptance and use of technology (UTAUT): Towards a revised theoretical model. *Information Systems Frontiers, 21*(3), 719–734.

Economides, N. (1996). The economics of networks. *International Journal of Industrial Organization, 14*(6), 673–699.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal, 50*(1), 25–32.

Eisenhardt, K. M., Graebner, M. E., & Sonenshein, S. (2016). Grand challenges and inductive methods: Rigor without rigor mortis (pp. 1113–1123).

Eisenmann, T., Parker, G., & Van Alstyne, M. (2011). Platform envelopment. *Strategic Management Journal, 32*(12), 1270–1285.

Ennen, E., & Richter, A. (2010). The whole is more than the sum of its parts—or is it? A review of the empirical literature on complementarities in organizations. *Journal of Management, 36*(1), 207–233.

Farrell, J., & Saloner, G. (1986). Installed base and compatibility: Innovation, product preannouncements, and predation. *The American Economic Review*, 940–955.

Finlay, S. (2017). *Artificial intelligence and machine learning for business*. Lancashire: Relativistic Books.

Franchini, M. (2013). The modern treatment of haemophilia: A narrative review. *Blood Transfusion, 11*(2), 178–182.

Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Boston: Pitman.

Freeman, R. E., Dmytriyev, S. D., & Phillips, R. A. (2021). Stakeholder theory and the resource-based view of the firm. *Journal of Management, 47*(7), 1757–1770.

Freeman, R. E., Harrison, J. S., & Wicks, A. C. (2007). *Managing for stakeholders: Survival, reputation and success*. New Heaven, C.T.: Yale University Press.

Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, H. (2021). The role of artificial intelligence and data network effects for creating user value. *Academy of Management Review, 46*(3), 534–551.

Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, H. (2022). Data network effects: Key conditions, shared data, and the data value duality. *Academy of Management Review, 47*(1), 189–192.

Gulati, R. (1999). Network location and learning: The influence of network resources and firm capabilities on alliance formation. *Strategic Management Journal, 20*(5), 397–420.

Gulati, R., Nohria, N., & Zaheer, A. (2000). Strategic networks. *Strategic Management Journal, 21*(3), 203–215.

Guo, H., Guo, A., & Ma, H. (2022). Inside the black box: How business model innovation contributes to digital start-up performance. *Journal of Innovation & Knowledge, 7*(2), Article 100188.

Giusti, J. D., Alberti, F. G., & Belfanti, F. (2020). Makers and clusters. Knowledge leaks in open innovation networks. *Journal of Innovation & Knowledge, 5*(1), 20–28. https://doi.org/10.1016/j.jik.2018.04.001

Godfrey, P. C. (2005). The relationship between corporate philanthropy and shareholder wealth: A risk management perspective. *Academy of Management Review, 30*(4), 777–798.

Hagiu, A. (2015). Strategic decisions for multisided platforms. *Top, 10*, 4–13.

Harting, T. R., Harmeling, S. S., & Venkataraman, S. (2006). Innovative stakeholder relations: When "ethics pays" (and when it doesn't). *Business Ethics Quarterly, 16*(1), 43–68.

Heugens, P. P., Van Den Bosch, F. A., & Van Riel, C. B. (2002). Stakeholder integration: Building mutually enforcing relationships. *Business & Society, 41*(1), 36–60.

Hillman, A. J., & Keim, G. D. (2001). Shareholder value, stakeholder management, and social issues: What's the bottom line? *Strategic Management Journal, 22*(2), 125–139.

Hilmersson, F. P., & Hilmersson, M. (2021). Networking to accelerate the pace of SME innovations. *Journal of Innovation & Knowledge, 6*(1), 43–49. 10.1016/j.jik.2020.10.001.

Horváth, D., & Szabó, R. Z. (2019). Driving forces and barriers of Industry 4.0: Do multinational and small and medium-sized companies have equal opportunities? *Technological Forecasting and Social Change, 146*, 119–132.

Hosmer, L. T., & Kiewitz, C. (2005). Organizational justice: A behavioral science concept with critical implications for business ethics and stakeholder theory. *Business Ethics Quarterly, 15*(1), 67–91.

Karmaker, C. L., Al Aziz, R., Palit, T., & Bari, A. M. (2022). Analyzing supply chain risk factors in the small and medium enterprises under fuzzy environment: Implications towards sustainability for emerging economies. *Sustainable Technology and Entrepreneurship*, Article 100032.

Katz, M. L., & Shapiro, C. (1985). Network externalities, competition, and compatibility. *The American Economic Review, 75*(3), 424–440.

Kela-Madar, N. (2021). Navigating through creativity and innovation paths: Waze navigation application case study. *Journal of Entrepreneurship Education, 24*(6), 1–12.

Kerr, D. (2013), July 25. Google reveals it spent $966 million in Waze acquisition. *CNET. CBS Interactive*. Retrieved Dec. 12, 2021.

Klein, H. K., & Myers, M. D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS Quarterly*, 67–93.

Kushwaha, A. K., Kar, A. K., & Dwivedi, Y. K. (2021). Applications of big data in emerging management disciplines: A literature review using text mining. *International Journal of Information Management Data Insights, 1*(2), Article 100017.

Lamiani, G., Bigi, S., Mancuso, M. E., Coppola, A., & Vegni, E. (2017). Applying a deliberation model to the analysis of consultations in haemophilia: Implications for doctor-patient communication. *Patient Education and Counseling, 100*(49), 690–695.

Liebowitz, S. J., & Margolis, S. E. (1994). Network externality: An uncommon tragedy. *Journal of Economic Perspectives, 8*(2), 133–150.

Lin, H., Lin, J., & Wang, F. (2022). An innovative machine learning model for supply chain management. *Journal of Innovation & Knowledge, 7*(4), Article 100276.

Locke, K. (2011). Filed research in management and organization studies. Reclaiming its tradition of discovery. *Academy of Management Annals, 5*, 613–652.

Lopes, T. J., Rios, R., Nogueira, T., & Mello, R. F. (2021). Prediction of hemophilia A severity using a small-input machine-learning framework. *NPJ Systems Biology and Applications, 7*(1), 1–8.

Macpherson, A., Jones, O., Zhang, M., & Wilson, A. (2003). Re-conceptualising learning spaces: Developing capabilities in a high-tech small firm. *Journal of Workplace Learning, 15*(6), 259–270.

Mayenberger, D. (2019). How to overcome modelling and model risk management challenges with artificial intelligence and machine learning. *Journal of Risk Management in Financial Institutions, 12*(3), 241–255.

Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research, 2*(3), 192–222.

Orlikowski, W. J., & Iacono, C. S. (2001). Research commentary: Desperately seeking the "IT" in IT research—A call to theorizing the IT artifact. *Information Systems Research, 12*(2), 121–134.

Park, C. (2020). Market entry strategies in a high-tech successive generations market: A case study of three semiconductor firms with different entry modes. *Journal of Business & Industrial Marketing*.

Parker, G. G., Van Alstyne, M. W., & Choudary, S. P. (2016). *Platform revolution: How networked markets are transforming the economy and how to make them work for you*. WW Norton & Company.

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. SAGE Publications Inc.

Pehrsson, A. (2009). Barriers to entry and market strategy: A literature review and a proposed model. *European Business Review, 21*(1), 64–77.

Perrini, F. (2006). SMEs and CSR theory: Evidence and implications from an Italian perspective. *Journal of Business Ethics, 67*, 305–316.

Porter, M. E. (1985). *Competitive advantage: Creating and sustaining superior performance*. New York: Free Press.

Rehm, S. V., Goel, L., & Junglas, I. (2017). Using information systems in innovation networks: Uncovering network resources. *Journal of the Association for Information Systems, 18*(8), 2.

Rochet, J.-C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association, 1*(4), 990–1029.

Roome, N., & Wijen, F. (2005). Stakeholder power and organizational learning. *Corporate Environmental Management, 10*, 69–76.

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development, 3*(3), 210–229.

Saura, J. R. (2021). Using data sciences in digital marketing: Framework, methods, and performance metrics. *Journal of Innovation & Knowledge, 6*(2), 92–102. 10.1016/j.jik.2020.08.001.

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review, 20*(3), 571–610.

Sheremata, W. A. (2004). Competing through innovation in network markets: Strategies for challengers. *Academy of Management Review, 29*(3), 359–377.

Shrestha, A., Eldar-Lissai, A., Hou, N., Lakdawalla, D. N., & Batt, K. (2017). Real-world resource use and costs of haemophilia A-related bleeding. *Haemophilia, 23*(4), e267–e275.

Smyth, S. M. (2019). The Facebook conundrum: Is it time to usher in a new era of regulation for big tech? *International Journal of Cyber Criminology, 13*(2), 578–595.

Spithoven, A., Vanhaverbeke, W., & Roijakkers, N. (2013). Open innovation practices in SMEs and large enterprises. *Small Business Economics, 41*(3), 537–562.

Statista (2021). https://www.statista.com/topics/3104/artificial-intelligence-ai-worldwide/ accessed: May 15th, 2021.

Sun, W., Zhao, Y., & Sun, L. (2020). Big data analytics for venture capital application: Towards innovation performance improvement. *International Journal of Information Management, 50*, 557–565.

Supyuenyong, V., Islam, N., & Kulkarni, U. (2009). Influence of SME characteristics on knowledge management processes: The case study of enterprise resource planning service providers. *Journal of Enterprise Information Management*.

Tagscherer, F., & Carbon, C. C. (2023). Leadership for successful digitalization: A literature review on companies' internal and external aspects of digitalization. *Sustainable Technology and Entrepreneurship*, Article 100039.

Tavory, I., & Timmermans, S. (2014). *2014 abductive analysis: Theorizing qualitative research*. Chicago: University of Chicago Press.

Terlaak, A., & Kim, J. Y. (2021). The paradoxical effect of prior operating experience and observational learning on new market entry. *Industrial and Corporate Change, 30*(4), 948–965.

Thietart, R. A. (2016). Strategy dynamics: Agency, path dependency, and self-organized emergence. *Strategic Management Journal, 37*(4), 774–792.

Thompson, R. L., Higgins, C. A., & Howell, J. M. (1991). Personal computing: Toward a conceptual model of utilization. *MIS Quarterly*, 125–143.

Towse, R. (2020). Dealing with digital: The economic organisation of streamed music. *Media, Culture & Society., 42*(7–8), 1461–1478.

Trapp, C. T., Kanbach, D. K., & Kraus, S. (2022). Sector coupling and business models towards sustainability: The case of the hydrogen vehicle industry. *Sustainable Technology and Entrepreneurship, 1*(2), Article 100014.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 425–478.

Watson, G. F., IV, Weaven, S., Perkins, H., Sardana, D., & Palmatier, R. W. (2018). International market entry strategies: Relational, digital, and hybrid approaches. *Journal of International Marketing, 26*(1), 30–60.

Weerakkody, V., Irani, Z., Kapoor, K., Sivarajah, U., & Dwivedi, Y. K. (2017). Open data and its usability: An empirical view from the Citizen's perspective. *Information Systems Frontiers, 19*(2), 285–300.

Wong, K. Y., & Aspinwall, E. (2005). An empirical study of the important factors for knowledge-management adoption in the SME sector. *Journal of Knowledge Management, 9*(3), 64–82.

World Bank (2020). https://www.worldbank.org/en/topic/smefinance (2020) Accessed: 03.08.2021.

Zhao, Y., Wen, S., Zhou, T., Liu, W., Yu, H., & Xu, H. (2022). Development and innovation of enterprise knowledge management strategies using big data neural networks technology. *Journal of Innovation & Knowledge, 7*(4), Article 100273.

Zimmerman, M. A., & Zeitz, G. J. (2002). Beyond survival: Achieving new venture growth by building legitimacy. *Academy of Management Review, 27*(3), 414–431.

Ricardo COSTA CLIMENT is a doctoral researcher and a PhD student in Information Systems at the Department of Informatics & Media, Uppsala University. He has studied various disciplines, including digital retail, omnichannel, law, and economics. His graduate studies have taken him to various universities in Spain, Austria, Ireland, Norway, Denmark, and Sweden. He holds bachelor's and master's degrees in Economics and Business in Law in Economic and Financial Management in University Research Training. He served as a researcher in the context of an EU-funded position within the Marie Sklodowska-Curie.

Dr. Darek M. Haftor is a professor of information systems and an expert in digital business models, productivity gains from digitalisation and the digital transformation of organisations. His research focuses on economic value creation from the use of digital technologies. Darek obtained a doctorate in Industrial Organisation at Chalmers University of Technology and has spent one and a half decades in the industry and the public sector.

Samuel Ribeiro Navarrete Samuel Ribeiro-Navarrete works at ESIC University (Spain) and is Visiting Professor PhD, University of Economics and Human Sciences in Warsaw (Poland). He has published papers in journals indexed in the WoS, such as International Journal of Information Management, Service Industries Journal, Service Business, Technological Forecasting & Social Change, International Journal of Entrepreneurial Behavior & Research, Journal of Business Research, Knowledge Management Research & Practice, among others. Editor of a Springer book. Currently is has been pointed as Editor of Economic Research (Q2 at SSCI-WoS).