

# Data-driven innovation: switching the perspective on Big Data

Data-driven  
innovation

Daniel Trabucchi and Tommaso Buganza  
*School of Management, Politecnico di Milano, Milan, Italy*

23

## Abstract

**Purpose** – The pervasive spread of digital technologies brought an incredible boost in data availability. Companies are dealing with massive amount of data that wait to be exploited. At the same time, scholars are providing different strategies and methods to help companies capture the value embedded in their data to foster innovation and improve the efficiency of existing processes. In these research studies, data are the by-product of something else, and they are a silent asset that needs to be exploited. What if data might be considered the final goal? The paper aims to discuss these issues.

**Design/methodology/approach** – The research is based on an exploratory multiple case study analysis, on the basis of three cases used as an illustration for new ideas. In particular, the gathered data are analyzed according to models previously presented in the literature review, building on and expanding them.

**Findings** – The research proposes a data-driven approach to innovation, offering a peculiar view of the innovation process. The trigger point is the need of data that let begin the entire development process of a complex system. In this perspective, the application that data are a by-product of the entire innovation process and not the primary output is peculiar since the vast majority of the literature consider data as the by-product of the primary product.

**Research limitations/implications** – Future research is needed to assess the replicability of the model outside the mobile app industry and to measure its performances. Nevertheless, this paper provides insights both for scholars and managers, enlarging the discussion on digital innovation and digital business models.

**Practical implications** – The results provide a development process to foster innovation relying on the need of data as the trigger point, guiding entrepreneurs and managers in the building process of the entire digital system.

**Originality/value** – Previous research studies often considered Big Data (BD) in innovation as a way to enlarge the current product offer or to make the innovation process more effective or efficient; this paper changes the perspective by considering BD as the trigger and the enabler of the entire digital innovation process.

**Keywords** Big Data, Internet of Things, Two-sided market, Innovation process, Digital innovation

**Paper type** Research paper

Received 29 January 2018

Revised 17 June 2018

Accepted 26 July 2018

## Introduction: assessing the role of digital technologies in innovation

In 2012, IBM published a research talking about 2.5 quintillion bytes of data created daily, which means roughly 0.9 ZB ( $10^{21}$ ) in a year. Estimations talk about 163 ZB ( $10^{21}$ ) a year, up from the 16.1 ZB registered in 2016 (IDC, 2017). These numbers are the result of the interconnected world we live in. Everything we do is enhancing the data creation process. Listening to a song is generating data on where, when and even how it has been listened to. Similarly, setting the temperature through our smartphone to go back in a warm and cozy house is generating data on our behavior, going out for a run is creating data on where we run, how fast we run, how many times we slow down and so on. This short list may be expanded to almost every single moment of our daily life, with a common characteristic that links all of them: our smartphone.

Smartphones are a bundle of sensors that can measure and gather a considerable amount of data that reach the cloud. Then, data are analyzed to give us back a service, to improve the service, may even be shared with other companies to improve other services and so on. This process creates new data at each step, leading to an enormous and continuous data chain production process.

These data are becoming so vast and relevant that both managers and scholars must consider them. Indeed, the opportunities enabled by Big Data (BD) have been largely studied in the recent literature (e.g. Buganza *et al.*, 2015; George and Lin, 2017) and remain a hot topic for research (e.g. Martinez and Walton, 2014). In particular, scholars also focused on



the strategies unveiled by BD to foster innovation (e.g. Trabucchi *et al.*, 2018), using them as a trigger to find meaningful directions to innovate within a company.

BD, Internet of Things and more generally all the technological developments that allow companies to gather a considerable amount and variety of data easily are part of the vast family of digital technologies, which are having a significant impact on innovation management.

Over the last years, we entered in the “golden age” of digital innovation (Fichman *et al.*, 2014), having a more pervasive and ubiquitous diffusion of digital technologies, which are having a significant impact on the innovation process. The term digital innovation, indeed, refers to the innovation process that relies on digital technologies to offer new products, services or processes (e.g. Martínez-López and Casillas, 2013; Nambisan *et al.*, 2017).

Scholars started considering the impact of these technologies on the innovation process and companies start considering them as new sources of competitive differentiation (Berman, 2012). Furthermore, scholars explore different fields and perspectives and the topics, such as their implementation process and related challenges (e.g. Gastaldi *et al.*, 2017) or their unveiled opportunities (e.g. Buganza *et al.*, 2015). Nevertheless, digital technologies have such a broad impact that makes them a particular complex and hot topic that still needs further development from a research perspective (e.g. Appio *et al.*, 2016; Brem *et al.*, 2016; Ganzaroli *et al.*, 2016; Rekonen and Björklund, 2016).

Nevertheless, this research aims to dig into the role of BD for the innovation process, showing a different perspective on the topic.

Nevertheless, if previous research studies (e.g. Johanson *et al.*, 2014; George and Lin, 2017) often considered these technologies as enabler for a more effective and efficient innovation process (e.g. increasing the ability to understand users’ needs), this paper aims to show how they can also play (another) a significant role in the innovation process: acting as the trigger factor of the entire innovation process.

On the basis of exploratory research based on relevant case studies as inspiration for new ideas, this paper digs in three companies that relied on digital technology in a very peculiar way, making their need for data – and the possibility to gather them – the enabling and starting factor of the entire innovation process.

First, the paper presents a literature review based on the role of BD phenomena in the innovation landscape. After the methodological section, the empirical results of the research are presented and furthered discussed through the theoretical lenses presented in the literature review.

In the end, implications for theory and managers are presented, highlighting how this research lets emerge a different view of the role of these digital technologies in the innovation process, considering them as the trigger instead of an enabler or an enhancer. Moreover, the suggested innovation process may be relevant for managers and entrepreneurs working in the digital environment, having a different perspective on how they may rely on the data they are gathering.

### **Literature review: the role of BD in innovation**

BD represents one of the most relevant emerging topics over the last years. Several definitions have been proposed over time (e.g. IBM, 2012; Brown *et al.*, 2011; Gandomi and Haider, 2015; De Mauro *et al.*, 2016), sharing the main features. In particular, the 3Vs model (McAfee and Brynjolfsson, 2012; Kaisler *et al.*, 2013; Anshari *et al.*, 2016) emerged from the others, highlighting the role of the following: volume (huge amount of data), velocity (continuous stream of data) and variety (different kinds of data from different sources) in defining BD. Later on, the Vs have been further expanded (e.g. Fan and Bifet, 2013; Kaisler *et al.*, 2013; Gandomi and Haider, 2015; Wamba-Fosso *et al.*, 2015; Del Vecchio *et al.*, 2018) with the concepts of veracity (using reliable data and confident interpretations), variability (managing and interpreting the continuous stream of data and their changes) and value (exploiting the value embedded in the data).

Different complementary trends have been considered relevant to enable and enhance the BD phenomenon (e.g. Dumbill, 2013). Among the others, the diffusion of smart devices (e.g. smartphones, tablets, and wearables) and social network platforms played a key role, since they are data generator (Lohr, 2012). Furthermore, the improvement and diffusion of cloud computing, giving the chance to have access and store data, boosted the diffusion significantly (Marston *et al.*, 2011), along with the cost reduction concerning computational power (Vajjhala and Ramollari, 2016).

Along with the diffusion of smartphones and digital services, other growing trends play a significant role in the BD diffusion. First, Internet of Things, which is a paradigm that can be considered the results of a convergence process of different visions over technology, integrating objects, connectivity and a semantic approach (Atzori *et al.*, 2010). It is commonly described “as a global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies” (ITU, 2012).

Human beings and their behaviors or opinion are often considered part of this evolving ecosystem, enabling even more potentialities regarding BD analytics (Chang, 2018). The integration of physical elements and people in generating data has been discussed in the concept of Internet of People (Conti *et al.*, 2017), defined as a complex socio-technical system where humans interact with their devices representing the key nodes of a network that creates valuable and exploitable data.

#### *The role of BD in the management literature*

The role of BD in the management literature is multifaceted. Initially, scholars focused mainly on their role to improve the decision-making process (McAfee and Brynjolfsson, 2012; Sahoo *et al.*, 2012). Companies started perceiving the chance to use business information and analytics as a critical variable to differentiate themselves within their industry (La Valle *et al.*, 2011). Soon after, different disciplines started paying attention to the phenomenon, studying potential applications in different fields, such as marketing (e.g. Chau and Xu, 2012; Leeftang *et al.*, 2014), customer engagement (Sahoo *et al.*, 2012; Xie *et al.*, 2016) or supply chain (Moretto *et al.*, 2017).

More in general, the concept of BD moved from a purely technological debate to a managerial one, seeing in the chance to create value through the transformation of such data assets as a key challenge (De Mauro *et al.*, 2016; Furtado *et al.*, 2017).

Later on, scholars pointed out the chance to leverage this kind of data to foster innovation, leveraging different data sources both internally and externally to the company (Sorescu, 2017).

One of the most significant examples is related to the chance to co-create with the customers, relying on historical customer data, gathering habits and preferences and so on (Furtado *et al.*, 2017).

Along with the previous mentioned technological opportunities – such as the chance to develop “smart services” able to gather data – this is pushing researcher to search for new comprehensive frameworks, merging the two traditional innovation paradigms: market-pull and technology-push (Geum *et al.*, 2016). Indeed, managers can leverage the growing diffusion of sensors and the decreasing barriers to access them to leverage entrepreneurship opportunities relying on different kind of data (Brown, 2017) and answering to different customers’ needs (e.g. Johanson *et al.*, 2014).

#### *The role of BD in innovation*

Recent research studies showed how BD could bring different kind of innovations, from process to product, from an architecture to a modular base (e.g. Caputo *et al.*, 2016; George and Lin, 2017). On the one hand, BD may help companies to improve the usage of

existing assets. Therefore, data can be used and re-used within the space where they have been developed fostering a so-called infrastructural innovation based on data (Tempini, 2017), or BD can be used to increase the level of personalization in services, leveraging data as a paradigm (e.g. Ng and Wakenshaw, 2017).

On the other, BD may help to extend company's capabilities in a new context and going beyond company's boundaries (Pellegrini, 2017). BD may drive an open innovation strategy to gain new business opportunities (Del Vecchio *et al.*, 2018) or even pursuing business model innovation to enlarge the operating domain of the company (Trabucchi *et al.*, 2018).

Scholars start also deepening in the organizational enablers that can facilitate this kind of innovation, such as customer centricity, data orientation management and the implementation of agile practices (Troilo *et al.*, 2017).

Even if the literature on the topic is growing quickly, there is still a lack of research on how BD can challenge the underlying mechanisms of value creation (Huberty, 2015).

Indeed, all the previous mentioned research studies considered BD as the by-product of some process. The focus so far has been on the strategies that can let companies unveil and capture this embedded value.

In particular, recent research showed how this by-product could be the sources of a new value appropriation mechanism leveraging a different kind of strategies. Three strategies to capture value through BD have been described: improving existing mechanisms such as advertising, understanding in a more profound way the current customers, or even finding a new kind of customer who can see a broader value in this data to search for it (Trabucchi *et al.*, 2017). At the basis of this model, there is the two-sided market theory. This kind of markets is characterized by the presence of two (or more) different groups of customer – the sides – that are linked together through a platform that internalizes the cross-side network externalities existing between the different groups (e.g. Rochet and Tirole, 2003).

These markets have been identified in different industries, such as credit cards or video consoles, and have been classified according to the type of transaction enabled by the platform (e.g. Rysman, 2009; Filistrucchi *et al.*, 2014).

Recent development on the topic showed how BD could challenge the traditional two-sided markets based on the advertising mechanisms (defined as a non-transaction two-sided market, since the platform is not allowing a direct transaction between the two sides). Ad-based businesses leverage a two-sided structure since a side (the advertisers) is looking forward to reach the first side (the readers or the viewers) through a Client-As-a-Target perspective to get their attention (Filistrucchi *et al.*, 2014; Trabucchi *et al.*, 2017).

The data gathered through a digital platform (e.g. a mobile app) can invert the fundamental relationship between the two sides, since the platform may leverage one of the above-mentioned strategies to engage a certain kind of customers on the second side, which show interest in the gathered data, leveraging a Client-As-a-Source perspective (Trabucchi *et al.*, 2017). It may be translated in different strategies: enhance advertising (matching a Client-As-a-Target and a Client-As-a-Source strategy to offer better-targeted advertising messages), e-Ethnography (which allows the app provider to gather data regarding specific behaviors of the end users) and data trading (which is based on the trading of data with third parties).

Researchers have recently started to point out the different kinds of innovation that can be pursued through BD. Chang (2018) pointed out the need for further research on how BD can provide incentives to foster innovation, especially while technological evolution is enabling the chance to gather specific data in a timely and cost-effective way (Gandomi and Haider, 2015). In the same perspective, Rindfleisch *et al.* (2017) showed how BD can both let companies exploit their value (innovation from data) as well as enable a particular kind of product innovation enabled by the chance to gather and analyze data (innovation as data). This research aims to deep in the second category, trying to show the chance to gather a particular kind of data as the main innovation outcome and to build around it the entire business model.

## Research method

A multiple case study method has been used in this research, due to the lack of previous studies in the field with a similar perspective and accordingly to the exploratory intent of the research (Yin, 2009).

The app industry is the empirical setting where the study took place for two main reasons. First, the pervasiveness of smartphones – with 2.87bn users expected by 2020 according to Statista (2016) – and the related continuous growth of available mobile apps make the field a source of an increasing number of meaningful innovations based on digital technologies. Second, as previously mentioned smartphones embed a number of sensors (e.g. compass, accelerometer, camera and GPS tracker) that allow users provider to gather a huge amount (volume) and kinds (variety) of data continuously (velocity) during the service delivery, making them a great example of Internet of Things and BD generators.

This research is inductive in nature, leveraging meaningful case studies as inspiration for new ideas, as suggested by Siggelkow (2007). Therefore, three cases are going to be presented and used for their inspiring power regarding the opportunities unveiled by BD and the models that managers and scholars can use to consider them with an innovation perspective. The chance to leverage on multiple sources of evidence and to rely on existing theories to read the gathered data increase the validity of our results (Gibbert *et al.*, 2008; Cook and Campbell, 1979). This research aims to propose a different perspective on the role of BD in innovation, without proposing a dogmatic model but stimulating new research studies and concrete managerial projects also leveraging this perspective on BD and digital technologies.

The data collection phase has been based on a qualitative web-based analysis, aiming to involve different kinds of sources. Firms' official websites have been used to have the institutional view on the service (how it is described, the kinds of data gathered, the kinds of feedbacks offered to the end users and the other players involved in the system). The official pages of the mobiles apps on the major stores (i.e. Google Play and the Apple App Store) were used to have a clear view on how the service is delivered to the end users and to have a look on the privacy policies of the services. Industry magazines and other publications have been used to have a deeper understanding of the underpinning mechanisms and the business models of the considered cases (searching mainly for statements from the managers as well as explanations on the dynamics of the system under observation).

The gathered documents have been analyzed through an iterative coding process, made up of three phases: reading, coding and interpreting (Saldaña, 2015). Based on the recommendations of Corbin and Strauss (2008), we used an open coding process, identifying key sentences from the documents and sorting them into first-order categories. Then, the categories have been combined through an axial coding process into higher-level categories by identifying the relationships between them and the analyzed literature. The gathered data have been analyzed through a comparison matrix of data, as suggested by Miles and Huberman (1984).

## Empirical results: inspiring innovation based on the power of Internet of Things

Three cases are going to be presented, focusing on how they leverage BD and reading them to the two-sided market models based on BD. The three cases are UpCoffee by Jawbone, Duolingo, and mPower: Mobile Parkinson Disease Study by Sage Bionetworks.

### *UpCoffee*

UpCoffe is a mobile app released by Jawbone in March 2014. In those years, Jawbone was getting massive popularity on the market for their smart wristband: UP by Jawbone. The wristband was announced in 2011, and it represented the first non-audio product sold by the company. This innovative activity tracker, based on a flexible rubber wristband, got a significant impact on the market, leveraging its several functions, from calories counting to activity monitoring and smart alarm, powered by the sleep monitoring (Stone, 2011).

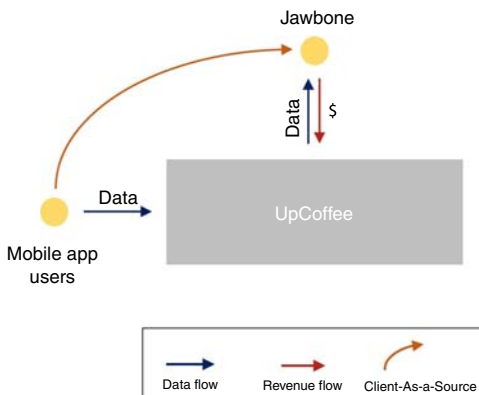
The system is based on a bundle of the tracker and the mobile apps that run on the smartphones, enabling data processing and direct feedback to the customer.

Interestingly, in 2014, the company released another app, which is the object of the case: UpCoffee. The underlying mechanism of the app is simple. Do you want to know how caffeine affects you? Use the app for seven days, and inform the app all the times you drink something with caffeine in it (selecting it from a comprehensive database that contains the information on the quantity of caffeine for the different drinks). In a week, the app will provide your – personal – relationship with caffeine (McCracken, 2014; Palladino, 2014). How? Considering the data that the wristband has been gathering – over time – monitoring the sleep and the data provided in the seven days.

According to Bandar Antabi – Head of special project at Jawbone when UpCoffee have been created UpCoffee – “The project started because we happened to be looking at which food items were logged the most often in our food system” (Dormehl, 2014), and after an internal hackathon convened by the internal data scientists, they understood they were missing some kind of data. When and how much caffeine do users consume? In order to get relevant insights on the relationship between caffeine and sleep patterns, the app has been developed, and results have been first given to the users and then published also on their blog (Raskin, 2014). The data gathered through this app enter in the Jawbone’s database, enhancing the service they can offer through the main app, improving the feedbacks and insights they offer to their end users through the main app (UP by Jawbone). A manager in the company said: “We want the hardware design, the software design, and the data to work in concert to help people achieve their goals” (Palladino, 2014).

According to the gathered data, the service seems to be based on a two-sided structure. The need for data triggered the system: Jawbone – working on UP by Jawbone – is searching for specific data (data consumption) to enhance their service. For this reason, they invest in a new service, UpCoffee, which has as customers the end users of their new app. Therefore, the two-sided structure is based, on the one hand, on the group of customers – represented by the end users of the mobile app – which receives the main service for free, providing the data necessary for the service delivered through the UpCoffee app. On the other hand, the company itself acts as a customer willing to have those data to enhance ongoing research and possibly to trade them somehow.

Figure 1 summarizes the case according to the two-sided market model (Trabucchi *et al.*, 2017): the app (the central rectangle) links the users with the overall company (which sustain the system economically), and the two sides are represented through the two dots. The blue (data) and red (money) arrows represent the exchanges between the sides and the service provider.



**Figure 1.**  
UpCoffee as a two-  
sided market with a  
Client-As-a-Source  
perspective

The two groups of customers are linked in a Client-As-a-Source (represented by the orange arrow), which means that the end users act as a source of value for the entire system receiving the service for free, relying on perspective with an e-ethnography perspective, aiming to get more information on the end users' behavior.

This strategy highlights how the data gathered serve as a way to enhance the non-transaction players (Jawbone in Figure 1) knowledge of users' behaviors, just like in applied ethnography but relying on high-tech devices.

Interestingly, it seems that the main goal of the entire system is not to provide an additional service to the wristband users or to engage them, but actually to gather those data.

### *Duolingo*

Duolingo is an American company founded by Luis von Ahn and Severin Hacker back in 2011. It is a free educational service that aims to spread the chance to learn a foreign language, offering more 65+ different courses.

The app is based on a personalization mechanism: "human tutors are scarce and expensive; we are using technology to give everyone access to personalized education in a scalable way" (Hof, 2015). Therefore, they rely on data coming from each single user to determine how the learning experience will continue; Laela Strudy, a partner at Google Capital which is investing in the project, stated: "Duolingo's mobile-first, adaptive, and gamified platform is changing the way people are learning languages across the world" (Hof, 2015).

Today the company's business model is expanding. It has direct revenues streams based on their test center – that allows the user to take a certification exam for \$49 staying at home – based on the potentialities of digital systems that allow them to check for the reliability of the results and avoid students to cheat. Moreover, the company is also working on other side projects, such as Tincards that allow users to improve their memory.

This case is particularly relevant for this study for two main reasons: the original aim of the project and the original business model.

Back in 2009, Luis von Ahn presented the reCAPTCHA project (later acquired by Google) which gave a meaning to the people all over the world who were solving captchas to enter in websites.

The main idea is to rely on the free work of millions of people who are trying to show to a system that they are human beings. For security reasons, we are asked to accomplish simple tasks that a computer would not be able to do, having then the chance to enter into the required system. The reCAPTCHA projects aim to create value through these activities, putting aside the basic CAPTCHA a second simple problem. At the beginning, it was used as the chance to read misspelled words to correct a mistake, while now it can also be used for different activities such as selecting pictures with a specific object. It aims to improve algorithms that are solving hard AI problems. This mechanism is now leveraged by Google to provide a free security service while gathering valuable data[1].

This project inspired Luis von Ahn, giving birth to something truly different, but based on a very similar mechanism: Duolingo.

Along with Severin Hacker, Luis Von Han aimed to have millions of people around the world translating information to enhance the value of the (mono language) knowledge on the web. In other words, they were searching for a way to have someone creating value for someone else, while they were doing something they wanted or needed to do. Before launching the app, he stated at TechCrunch: "The solution was to transform language translation into something that millions of people want to do, and that helps with the problem of lack of bilinguals: language education. It is estimated that there are over 1 billion people learning a foreign language. So, the site that we've been working on, Duolingo, will be a 100% free language learning site in which people learn by helping to translate the Web. That is, they learn by doing" (Sieglar, 2011). A few years after, they completed the original

business model involving companies like CNN and BuzzFeed and offering them real-time translations. Von Ahn stated on the company blog: “As you know, Duolingo is committed to providing free language education for the world. From the beginning, our plan to finance the platform has been to have our students translate real-world documents as they practice their foreign language skills. We like this model because it’s a mutually beneficial arrangement: students receive high-quality, completely free language education, and organizations get translation services powered by the students. Today’s announcement that two major publishers are financing our operation by translating their content with us is a significant milestone in keeping Duolingo free forever” (von Ahn, 2013).

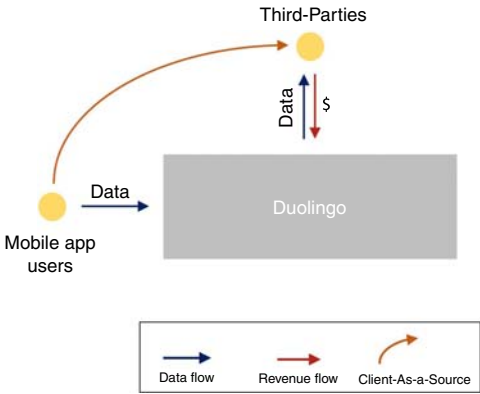
According to the case description, a two-sided structure emerges. The willingness to have a system to translate sentences from a language to another is the trigger point, along with the chance to realize that people who search for a (free or cheap) way to learn a language may be the right persons to do these activities when a certain level is reached. This leads to the identification of the two groups of customers. On the first side, there are end users receiving a free service (education) while offering to the service provider different kinds of data, among them translations from the language they are studying in their language. On the other side, companies are paying to get those translations which have been described as the trigger point of the entire system.

Figure 2 summarizes the overall mechanisms, having the mobile app users who receive a free service offering back translations in their language (data) that are then sold to the second side players (such as CNN and BuzzFeed in the example) in a Client-As-a-Source perspective, leveraging a data trading strategy (Trabucchi *et al.*, 2017). In this case, the second side (non-transaction, the third parties) is outside the firm and there is an economic transaction with the main service provider (Duolingo itself).

*mPower: Mobile Parkinson Disease Study*

In April 2015, Sage Bionetworks, a not-for-profit research organization launched Parkinson mPower study app.

The goal of the project is ambitious, noble and extremely clear. Researchers are searching for new ways to manage the symptoms of Parkinson’s disease. Sage Bionetworks found a way to collect valuable data on the symptoms and their evolution. The final goal is to understand the reasons why different people have different symptoms and the variation of side effects over time as well as understand how to help people managing the different symptoms in everyday life, as they declare on their official website[2].



**Figure 2.**  
Duolingo as a two-  
sided market with a  
Client-As-a-Source  
perspective



The user experience is based on a mix of surveys and tasks based on the different sensors of the smartphones (e.g. touchscreen, microphone and camera), allowing the gathering of a considerable amount of data. The final goal is to get valuable data for research, to enhance the knowledge on the diseases and to improve the quality of life for people with Parkinson's disease. The mobile app can be used both by people affected and not affected by the disease, allowing the company to have age-matched control data. The users have access to a dashboard that can help to keep track of the health data.

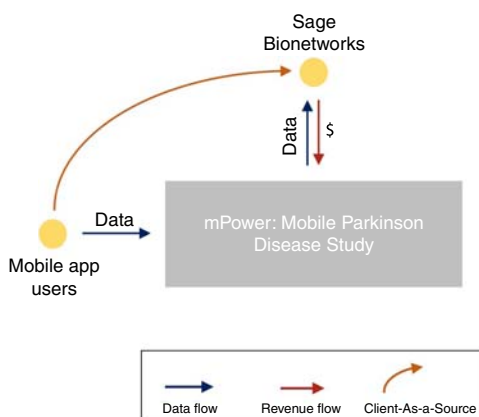
This mobile app brought a significant result, relying on data from more than 9,500 people enrolled in the study and providing millions of data point collected on a continuous basis through the smartphones involved (Moore, 2016). The data gathered through the app were brought to the publication of scientific papers on the topic (e.g. Bot *et al.*, 2016). Paul Tarini a Senior Program Officer at the Robert Wood Johnson Foundation said "This ground-breaking dataset is a wonderful example of how these new technologies and platforms can help us build a robust culture of health. They make it easier for more people and researchers to participate in efforts to understand diseases and how to manage them and, in turn, all of us benefit" (Moore, 2016).

Also in this third case, the need and the search for data is the trigger point of the entire system. Research studies – and in particular Sage Bionetworks – search for precious data on how diseases evolve. An app offering something in exchange – the chance to track the health data – has been created and delivered to those willing to participate in the study. Through the information gathered, also in this third case, a two-sided structure emerges. The first side is represented by the end users of the mobile app, who receive a service through the exercises provided by the app, while giving back to the system the usage data. These data are then used by the second side player – the company itself – to enhance their research studies on the disease object of the app.

Figure 3 summarizes the two-sided model, showing the mobile app users as the data providers in a Client-As-a-Source perspective who are linked with the second side, represented by the company itself that can (according to the decision of the single user) allow the access to other researchers in an e-ethnography perspective (Trabucchi *et al.*, 2017), also in this case having a way to observe the people behaviors and having direct information on them.

### Cross-cases analysis

This brief section aims to provide an overall view on the three cases in order to sum up the main similarities and differences previously presented, before entering in the discussion of the findings under the lenses of previous literature in the field.



**Figure 3.**  
mPower: Mobile  
Parkinson Disease  
Study as a two-sided  
market with a Client-  
As-a-Source  
perspective

Table I summarizes the main dimensions relevant to this study. First, the two different groups of customers representing the basic elements of a two-sided market, along with their needs solved by the app provider. Second, the rationale for having a Client-As-a-Source relationship (unidirectional cross-side network externalities). Third, the kind of Client-As-a-Source strategy implemented in the system. Finally, the trigger point of the innovation process is pointed out.

**Discussion: a development model to foster innovation starting from data**

*From data as by-product to data as the product*

The above presented case studies share a common structure: they are based on a non-transaction two-sided market leveraging a Client-As-a-Source perspective (Trabucchi *et al.*, 2017) relying on the data they gather through the service delivered to the end users. The peculiarity that links the three illustrative cases is the trigger point of the entire system: the need of data.

Ongoing research studies in the innovation field tend to show how products and services are creating data as by-product. In this perspective, the entire streams on the Internet of Things (e.g. Atzori *et al.*, 2010; Buganza *et al.*, 2015) and on the Internet of People (e.g. Conti *et al.*, 2017; Chang, 2018) explore how digital products and services generate and use a number of kinds of data to deliver the service to the end users. Leveraging on these streams, several research works show how these data can be used to unveil innovation opportunities (e.g. Geum *et al.*, 2016; Brown, 2017; Caputo *et al.*, 2016; George and Lin, 2017), exploiting their value (e.g. Del Vecchio *et al.*, 2018; Trabucchi *et al.*, 2018). This research aims to show the other side of the same coin: can these data be something different than the by-product of a product service system?

	First side	Second side	(Unidirectional) Cross-side network externalities	Kind of Client-As-a-Source strategy	Trigger points
UpCoffee	End users Need: searching for information on their reaction to caffeine	Jawbone Need: to gather data on coffee consumption to enlarge the current data set	An increase in the numbers of end users bring to an higher value for Jawbone, which has a larger data set to use to increase the level of personalization and recommendation of the service	e-Ethnography	Missing data in their data sets to increase the suggestions and the recommendations to UP by Jawbone users
Duolingo	End users Need: learning a foreign language	Third parties Need: to have translations of short sentences (mainly from various languages to English)	An increase in the number of end users bring to an higher value for third parties which may have faster translations (and even from a higher number of languages)	Data trading	A system to translate sentences from one language to another, while translators are doing something different (learning)
mPower	End users Need: to keep track to health data (and also to contribute to the research)	Sage Bionetworks Need: to have data on how disease evolves to research on it	An increase in the number of end users brings to a higher value for Sage Bionetworks increasing the statistical relevance of the data gathered	e-Ethnography	Lack of data on how diseases evolve to research on them

**Table I.**  
Cross-cases analysis

According to the findings of this paper, companies can leverage the opportunities provided by the BD and the pervasive diffusion of sensors to gather a specific kind of data and therefore to design an entire system and business model around the need of data.

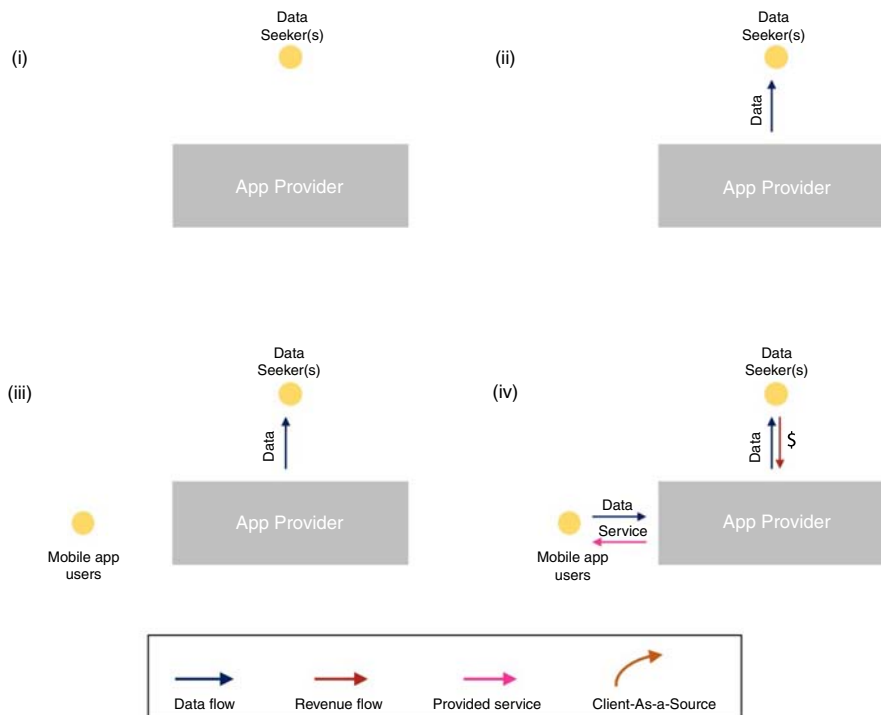
Jawbone enlarged the service offer with another side app to gather a specific kind of data, such as caffeine consumption. Duolingo designed an entire learning experience to bring together two contrasting needs: the search for free language education and the willingness to translate mono language knowledge. mPower designed a system to supply researchers with the data they need to carry on research.

Three different fields – activity tracking, education and research on medical diseases – have different overall business models – a side project, a joint system matching needs and a not-for-profit organization – but only one model fosters innovation through data. End users are considered a valuable source to provide meaningful data that can be analyzed, cross-checked, integrated and aggregated to provide knowledge to a third player in the system.

#### *A development process for non-transaction two-sided markets*

The insights derived from these illustrative cases let emerge a design model for non-transaction two-sided markets (Filistrucchi *et al.*, 2014) based on a Client-As-a-Source perspective (Trabucchi *et al.*, 2017).

The first step is based on the identification of the innovation issue that data can solve (Figure 4). Companies can have internal need of data – such as in the case of Jawbone, searching for caffeine consumption data to enlarge the analyses they were already performing on the previously gathered data – as well as identifying the need of external players – such as Sage Bionetworks, realizing the need to have more spread and precise data



**Figure 4.**  
A data-driven approach to innovation through two-sided markets

on how a certain diseases actually evolve in order to enhance their research works in the field. The two constraints of the innovation issue are: the chance to solve the need with identifiable kinds of data; the chance to retrieve those data through sensors embedded in smartphones and related smart objects.

The second phase is based on the deep understanding of the kind of data needed to solve the innovation issue (Figure 4), understanding what kind of data and analyses are needed to transform gatherable data into valuable knowledge for the data seeker(s). For example, Jawbone understood the need to have precise information on caffeine consumption over the day that could be matched with the data they already had (e.g. sleep patterns and physical activities).

The third phase is based on the involvement of the second side of the two-sided market (Figure 4). Who can provide the needed data? Who can be the end users who may create and supply those data? This is the phase that directly leads to the following one, defining the end users target of the mobile app. For example, Duolingo in the phase realized that the people who may translate short sentences from the web were actually the same people who were learning a foreign language, using the translation in their mother tongue as a learning exercise (and having the chance to later cross-check the translation with the ones provided by other users to increase the reliability).

The final phase (Figure 4) is based on the overall design of the system. First of all, it includes the service design, defining the kind of interaction and user experience of the mobile app. The requirements of these phase are mainly two: the definition of a meaningful experience for the end users that is going to create and supply the data and the definition of an engagement mechanism that allows the system to sustain data creation process, at least for the minimum necessary time span. Since two-sided markets are based on a double value proposition (Muzellec *et al.*, 2015), in this phase it is important to understand which is the value proposition that is going to be presented to the end users. For example, Jawbone understood the value for users to have a personalized feedback on the personal relationship between caffeine consumption and the sleep patterns; this leads to the creation of the overall user experience of the mobile app. Similarly, mPower realizes that the chance to offer a dashboard giving the chance to track health data is meaningful for end users and Duolingo designed an entire language service leveraging on the system they had in mind (relying on the previous experience of reCAPTCHA).

Finally, this phase brings to the definition of the complete business model, defining the revenue streams that may or may not involve the end users and that may receive the service for free, as in the mentioned cases or partially pay to sustain the entire system in a freemium perspective (Trabucchi *et al.*, 2017).

#### *Data-driven innovation under the lenses of previous literature*

As previously mentioned, scholars recently showed how BD could challenge the traditional innovation approaches, often described as a dyadic choice between a market-pull and a technology-push, mixing the two perspectives (Geum *et al.*, 2016). Our analysis builds on the same perspective, showing how companies can leverage technological opportunities provided by smartphones and digital devices integrating a market-pull perspective in the game.

Indeed, if previous studies focused mainly on how to exploit the value of the gathered data (Del Vecchio *et al.*, 2018), seeing customer centricity as a driver to enhance this process (Troilo *et al.*, 2017) this research asks for a different standpoint. The customer should still be at the center of the design process. Nevertheless, the customer we are referring to is not the end user.

Relying on the concept of two-sided market (Rochet and Tirole, 2003), this research builds on previous attempts to highlight two different groups of players involved in the value creation and value capturing process related to BD (Trabucchi *et al.*, 2017). The shift provided by this study is to understand where can be found the customers' need that triggers the innovation process, leveraging those literature streams that stress the ability of BD to change the underlying mechanisms of value creation (Huberty, 2015).

Leveraging the non-transaction model of two-sided markets (Rochet and Tirole, 2003; Filistrucchi *et al.*, 2014), our research shows how the service provider may involve more than one group of customers. Previous research studies in the field showed the chance to create a non-transaction two-sided market based on the opportunities provided by BD as a way to create a sustainable business model leveraging on innovative value appropriation strategies (Trabucchi *et al.*, 2017). This research pushes the model to another step, considering customers who are searching for data as the starting point of the entire system, changing the perspective on the role of data.

Previous research studies considered BD as a by-product of a system that may be a valuable asset to be exploited (e.g. Buganza *et al.*, 2015), recently Rindfleisch *et al.* (2017) pointed out how data can even enable a particular kind of product innovation, while this research shows how the BD can take the stand at the very beginning of the innovation process, representing the final goal to be reached at the end of the process, moving the perspective on the digital application that allows the gathering: from the central service to a by-product useful to gather data.

### Conclusion and implications

This research aimed to offer a new perspective on the role that BD may have in the innovation process. Starting from the literature streams that show how BD and related digital opportunities may be a good way to enhance the innovation process exploiting the value of those data that have been gathered as a by-product during the main service offered to the end users.

Based on the received literature on two-sided markets, this research is rooted in those market structures that involve two different groups of players – linked through specific kinds of cross-side network externalities (Katz and Shapiro, 1985; Rochet and Tirole, 2003) – creating a system based on the interaction between the two sides and the service provider. In particular, this research is based on the model that explains how the role of the end users may shift from a passive target to an active source of value for the system, thanks to the possibilities enabled by digital technologies (Trabucchi *et al.*, 2017).

The findings present a different approach to the innovation process based on the role of BD. In particular, the cases show companies that started with the need for certain kinds of data as the trigger of the innovation process.

This observation contributes to the literature in two different ways: first contributing to the ongoing debate on the role of BD and more broadly digital technologies in innovation and second to the growing literature on two-sided markets.

This research contributes to the ongoing debate on the opportunities unveiled by digital technologies and related data (e.g. Gandomi and Haider, 2015; Chang, 2018), building in particular on the growing streams that talk about “innovation as data” (Rindfleisch *et al.*, 2017).

Previous research tends to show how data are a by-product to be exploited (Ng and Wakenshaw, 2017; Tempini, 2017; Del Vecchio *et al.*, 2018; Trabucchi *et al.*, 2018), while in this case data are the primary output of the innovation process.

This can be directly linked to the second contribution, looking at the literature on two-sided markets. Non-transaction two-sided markets have been traditionally considered as a design choice regarding the business model to profit from services offered for free or almost for free to the end users, both in the Client-As-a-Target (Filistrucchi *et al.*, 2014) and in the Client-As-a-Source perspectives (Trabucchi *et al.*, 2018). Which means deciding to add the second (non-transaction) side to monetize what is being offered to the first side (end users).

The results of our research are showing a different mechanism: the chance to rely on a two-sided structure is an intrinsic characteristic of the entire system, which is a common feature for transaction two-sided markets (Muzellec *et al.*, 2015; Tauscher and Laudien, 2018). Still these kinds of two-sided markets are still keeping the traditional characteristics

of this market structure, such as the need of a double value proposition (Muzellec *et al.*, 2015) as presented in the cross-cases analysis table, and the chance to subsidize one of the two sides (Parker and Van Alstyne, 2005). Still, this research contributes to propose a different view and suggests a process for those non-transaction two-sided markets based on the usage of data and a Client-As-a-Source strategy (Trabucchi *et al.*, 2017).

From a managerial perspective, the contribution is straightforward. First, the chance to think about the opportunities unveiled by digital technologies in a new way. In particular, managers and entrepreneurs should start thinking about BD as not only something that can be generated on purpose to answer to specific innovation problems but also as a by-product to be exploited.

Second, the proposed model offers a first guide in the designing of such a system. In particular, it highlights the need to highlight the problem to be solved thorough gatherable data, the chance to clearly identify those data and the end users who may generate them to close the design process thinking about the value proposition that can be offered to the end users to ask them to generate the searched data.

This is changing the perspective from data as a by-product to data as the primary asset of the entire system designed around a digital service.

In this perspective, it is essential to discuss the role of the end users in the system. If traditional non-transaction two-sided markets based on advertising messages were giving to the end users the passive role of a target, in this perspective end users play an active and precious role in the system. End users are the main source of value creation in the system, creating and supplying data while enjoying (maybe for free) a service [...] which is not necessary the fundamental reason why the entire system has been created.

This research does not mean to be exhaustive, having an exploratory internet. Indeed, it is based on a small and convenient sample, chosen for its illustrative power (Siggelkow, 2007). Furthermore, it is based on secondary sources, which have been considered adequate for the purpose of the research (showing the existence of different strategies to rely on BD to foster the innovation process), but which represent a limitation of the research. Furthermore, the cases are all based on the mobile app industry, which is coherent with the setting of the research, but which decrease the chance to generalize the findings.

The limitations mentioned above allow several avenues of further research. The chance to enlarge the research to primary sources is a possible development. In particular, the chance to answer to more specific questions (e.g. How can companies rely on the data-driven innovation process? Which are the opportunities and challenges related to this kind of innovation approach?) would be enabled, building on the results of this paper. Indeed, the phases presented in this research may be further explored through different methodologies (e.g. qualitative surveys with similar companies) or also in different empirical fields (e.g. smart products).

## Notes

1. Further information at [www.google.com/recaptcha/intro/android.html](http://www.google.com/recaptcha/intro/android.html)
2. For further information visit: <https://parkinsonmpower.org>

## References

- Anshari, M., Alas, Y. and Guan, L.S. (2016), "Developing online learning resources: Big Data, social networks, and cloud computing to support pervasive knowledge", *Education and Information Technologies*, Vol. 21 No. 6, pp. 1663-1677.
- Appio, F.P., Martini, A., Massa, S. and Testa, S. (2016), "Unveiling the intellectual origins of social media-based innovation: insights from a bibliometric approach", *Scientometrics*, Vol. 108 No. 1, pp. 355-388.

- 
- Atzori, L., Iera, A. and Morabito, G. (2010), "The Internet of Things: a survey", *Computer Networks*, Vol. 54 No. 15, pp. 2787-2805.
- Berman, S.J. (2012), "Digital transformation: opportunities to create new business models", *Strategy & Leadership*, Vol. 40 No. 2, pp. 16-24.
- Bot, B.M., *et al.* (2016), "The mPower study, Parkinson disease mobile data collected using Research Kit", *Scientific Data*, Vol. 3, doi: 10.1038/sdata.2016.11.
- Brem, A., Maier, M. and Wimschneider, C. (2016), "Competitive advantage through innovation: the case of Nespresso", *European Journal of Innovation Management*, Vol. 19 No. 1, pp. 133-148.
- Brown, B., Chui, M. and Manyika, J. (2011), "Are you ready for the era of 'Big Data'", *McKinsey Quarterly*, Vol. 4 No. 1, pp. 24-35.
- Brown, T.E. (2017), "Sensor-based entrepreneurship: a framework for developing new products and services", *Business Horizons*, Vol. 60 No. 6, pp. 819-830.
- Buganza, T., Dell'Era, C., Pellizzoni, E., Trabucchi, D. and Verganti, R. (2015), "Unveiling the potentialities provided by new technologies: a process to pursue technology epiphanies in the smartphone app industry", *Creativity and Innovation Management*, Vol. 24 No. 3, pp. 391-414.
- Caputo, A., Marzi, G. and Pellegrini, M.M. (2016), "The Internet of Things in manufacturing innovation processes: development and application of a conceptual framework", *Business Process Management Journal*, Vol. 22 No. 2, pp. 383-402.
- Chang, V. (2018), "A proposed social network analysis platform for Big Data analytics", *Technological Forecasting and Social Change*, No. 130, pp. 57-68.
- Chau, M. and Xu, J. (2012), "Business intelligence in blogs: understanding consumer interactions and communities", *MIS Quarterly*, Vol. 36 No. 4, pp. 1189-1216.
- Conti, M., Passarella, A. and Das, S.K. (2017), "The Internet of People (IoP): a new wave in pervasive mobile computing", *Pervasive and Mobile Computing*, Vol. 41, pp. 1-27.
- Cook, T. and Campbell, D. (1979), *Quasi-experimentation: Design and Analysis for Field Settings*, Rand-McNally, New York, NY.
- Corbin, J. and Strauss, A. (2008), *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, 3rd ed., Sage Publications, Los Angeles, CA.
- De Mauro, A., Greco, M. and Grimaldi, M. (2016), "A formal definition of Big Data based on its essential features", *Library Review*, Vol. 65 No. 3, pp. 122-135.
- Del Vecchio, P., Di Minin, A., Petruzzelli, A.M., Panniello, U. and Pirri, S. (2018), "Big Data for open innovation in SMEs and large corporations: trends, opportunities, and challenges", *Creativity and Innovation Management*, Vol. 27 No. 1, pp. 6-22.
- Dormehl, L. (2014), "Coffee app lets caffeine junkies make sense of their habit", *Cult of Mac*, July 18 [www.cultofmac.com/287867/jawbone-wants-quantify-caffeine-consumption-habits/](http://www.cultofmac.com/287867/jawbone-wants-quantify-caffeine-consumption-habits/) (accessed November 25, 2017).
- Dumbill, E. (2013), "Making sense of big data", *Big Data*, Vol. 1 No. 1, pp. 1-2.
- Fan, W. and Bifet, A. (2013), "Mining Big Data: current status, and forecast to the future", *ACM SIGKDD Explorations Newsletter*, Vol. 14 No. 2, pp. 1-5.
- Fichman, R.G., Dos Santos, B.L. and Zheng, Z.E. (2014), "Digital innovation as a fundamental and powerful concept in the information systems curriculum", *MIS Quarterly*, Vol. 38 No. 2, pp. 329-345.
- Filistrucchi, L., Geradin, D., Van Damme, E. and Affeldt, P. (2014), "Market definition in two-sided markets: theory and practice", *Journal of Competition Law and Economics*, Vol. 10 No. 2, pp. 293-339.
- Furtado, L., Dutra, M. and Macedo, D. (2017), "Value creation in Big Data scenarios: a literature survey", *Journal of Industrial Integration and Management*, Vol. 2 No. 1, pp. 1750002-1-1750002-17.

- Gandomi, A. and Haider, M. (2015), "Beyond the hype: Big Data concepts, methods, and analytics", *International Journal of Information Management*, Vol. 35 No. 2, pp. 137-144.
- Ganzaroli, A., De Noni, I., Orsi, L. and Belussi, F. (2016), "The combined effect of technological relatedness and knowledge utilization on explorative and exploitative invention performance post-M & A", *European Journal of Innovation Management*, Vol. 19 No. 2, pp. 167-188.
- Gastaldi, L., Pietrosi, A., Lessanibahri, S., Paparella, M., Scaccianoce, A., Provenza, G., Corso, M. and Gridelli, B. (2018), "Measuring the maturity of business intelligence in healthcare: supporting the development of a roadmap toward precision medicine within ISMETT hospital", *Technological Forecasting and Social Change*, Vol. 128, pp. 84-103.
- George, G. and Lin, Y. (2017), "Analytics, innovation, and organizational adaptation", *Innovation: Management, Policy and Practice*, Vol. 19 No. 1, pp. 16-22.
- Geum, Y., Jeon, H. and Lee, H. (2016), "Developing new smart services using integrated morphological analysis: integration of the market-pull and technology-push approach", *Service Business*, Vol. 10 No. 3, pp. 531-555.
- Gibbert, M., Ruigrok, W. and Wicki, B. (2008), "What passes as a rigorous case study?", *Strategic Management Journal*, Vol. 29 No. 13, pp. 1465-1474.
- Hof, R. (2015), "AI-Powered Language learning app Duolingo Nets \$45 million funding", *Forbes*, June 10 [www.forbes.com/sites/roberthof/2015/06/10/ai-powered-language-learning-app-duolingo-nets-45-million-funding/#7b59232d2bd7](http://www.forbes.com/sites/roberthof/2015/06/10/ai-powered-language-learning-app-duolingo-nets-45-million-funding/#7b59232d2bd7) (accessed November 25, 2017).
- Huberty, M. (2015), "Awaiting the second Big Data revolution: from digital noise to value creation", *Journal of Industry, Competition and Trade*, Vol. 15 No. 1, pp. 35-47, doi: 10.1007/s10842-014-0190-4.
- IBM (2012), "What is Big Data?", available at: [www-01.ibm.com/software/data/bigdata/](http://www-01.ibm.com/software/data/bigdata/) (accessed November 25, 2017).
- IDC (2017), "Data age 2025: the evolution of data to life-critical", available at: [www.seagate.com/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf](http://www.seagate.com/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf)
- ITU (2012), "Overview of the Internet of Things", Series Y: Global Information Infrastructure, Internet Protocol Aspects and Next-Generation Networks, International Telecommunication Union, ITU-T Y.2060, Geneva, June, available at: [www.itu.int/rec/T-REC-Y.2060-201206-I](http://www.itu.int/rec/T-REC-Y.2060-201206-I)
- Johanson, M., Belenki, S., Jalminger, J., Fant, M. and Gjertz, M. (2014), "Big automotive data: leveraging large volumes of data for knowledge-driven product development", *2014 IEEE International Conference on Big Data*, ISBN: 978-1-4799-5666-1, Washington, DC, pp. 736-741.
- Kaisler, S., Armour, F., Espinosa, J.A. and Money, W. (2013), "Big Data: issues and challenges moving forward", paper presented at the 2013 46th Hawaii International Conference on System Sciences (HICSS), Wailea, Maui, HI, January 7-10, pp. 995-1004.
- Katz, M.L. and Shapiro, C. (1985), "Network externalities, competition, and compatibility", *The American Economic Review*, Vol. 75 No. 3, pp. 424-440.
- La Valle, S., Lesser, E., Shockley, R., Hopkins, M.S. and Kruschwitz, N. (2011), "Big Data, analytics and the path from insights to value", *MIT Sloan Management Review*, Vol. 52 No. 2, p. 21.
- Leefflang, P.S., Verhoef, P.C., Dahlström, P. and Freundt, T. (2014), "Challenges and solutions for marketing in a digital era", *European Management Journal*, Vol. 32 No. 1, pp. 1-12.
- Lohr, R. (2012), "The age of Big Data", *The New York Times*, February 11, available at: [www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html](http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html) (accessed December 18, 2016).
- McAfee, A. and Brynjolfsson, E. (2012), "Big Data: the management revolution", *Harvard Business Review*, Vol. 90 No. 10, pp. 61-67.
- Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J. and Ghalsasi, A. (2011), "Cloud computing – the business perspective", *Decision Support Systems*, Vol. 51 No. 1, pp. 176-189.
- Martinez, M.G. and Walton, B. (2014), "The wisdom of crowds: the potential of online communities as a tool for data analysis", *Technovation*, Vol. 34 No. 4, pp. 203-214.



- Martínez-López, F.J. and Casillas, J. (2013), "Artificial intelligence-based systems applied in industrial marketing: an historical overview, current and future insights", *Industrial Marketing Management*, Vol. 42 No. 4, pp. 489-495.
- Miles, M.B. and Huberman, A.M. (1984), "Drawing valid meaning from qualitative data: toward a shared craft", *Educational Researcher*, Vol. 13 No. 5, pp. 20-30.
- Moore, C. (2016), "6-month data from Parkinson's mPower App Study in 9,500 people released to researcher's", *Parkinson's News Today*, March 7, available at: <https://parkinsonsnewstoday.com/2016/03/07/sage-bionetworks-releases-first-of-its-kind-data-from-parkinsons-iphone-study/> (accessed November 25, 2017).
- Moretto, A., Ronchi, S. and Patrucco, A.S. (2017), "Increasing the effectiveness of procurement decisions: the value of Big Data in the procurement process", *International Journal of RF Technologies*, Vol. 8 No. 3, pp. 79-103.
- Muzellec, L., Ronteau, S. and Lambkin, M. (2015), "Two-sided internet platforms: a business model lifecycle perspective", *Industrial Marketing Management*, Vol. 45, pp. 139-150.
- Nambisan, S., Lyytinen, K., Majchrzak, A. and Song, M. (2017), "Digital innovation management: reinventing innovation management research in a digital world", *MIS Quarterly*, Vol. 41 No. 1, pp. 223-238.
- Ng, I.C.L. and Wakenshaw, S.Y.L. (2017), "The Internet of Things: review and research directions", *International Journal of Research in Marketing*, Vol. 34 No. 1, pp. 3-21.
- McCracken, H. (2014), "Attention caffeine addicts: Jawbone's Up Coffee App is for you", *Technolizer*, March 6, available at: <http://time.com/13683/attention-caffeine-addicts-jawbones-up-coffee-app-is-for-you/> (accessed November 25, 2017).
- Palladino, V. (2014), "Up Coffee is Jawbone's solution for America's caffeine addiction", *The Verge*, March 6, available at: [www.theverge.com/2014/3/6/5474738/jawbone-up-coffee-app-tracks-caffeine-consumption-to-help-you-sleep-better](http://www.theverge.com/2014/3/6/5474738/jawbone-up-coffee-app-tracks-caffeine-consumption-to-help-you-sleep-better) (accessed November 25, 2017).
- Parker, G.G. and Van Alstyne, M.W. (2005), "Two-sided network effects: a theory of information product design", *Management Science*, Vol. 51 No. 10, pp. 1494-1504.
- Pellegrini, T. (2017), "Semantic metadata in the publishing industry – technological achievements and economic implications", *Electronic Markets*, Vol. 27 No. 1, pp. 9-20.
- Raskin, A. (2014), "U Coffee: when brains brew", The Jawbone Blog, June 18, available at: <https://jawbone.com/blog/up-coffee/> (accessed November 25, 2017).
- Rekonen, S. and Björklund, T.A. (2016), "Adapting to the changing needs of managing innovative projects", *European Journal of Innovation Management*, Vol. 19 No. 1, pp. 111-132.
- Rindfleisch, A., O'Hern, M. and Sachdev, V. (2017), "The digital revolution, 3D printing, and innovation as data", *Journal of Product Innovation Management*, Vol. 34 No. 5, pp. 681-690.
- Rochet, J. and Tirole, J. (2003), "Platform competition in two-sided markets", *Journal of the European Economic Association*, Vol. 1 No. 4, pp. 990-1029.
- Rysman, M. (2009), "The economics of two-sided markets", *The Journal of Economic Perspectives*, Vol. 23 No. 3, pp. 125-143.
- Sahoo, N., Singh, P.V. and Mukhopadhyay, T. (2012), "A hidden Markov model for collaborative filtering", *MIS Quarterly*, Vol. 36 No. 4, pp. 1329-1356.
- Saldaña, J. (2015), *The Coding Manual for Qualitative Researchers*, Sage, Thousand Oaks, CA.
- Siegler, M.G. (2011), "Meet Duolingo, Google's next acquisition target; learn a language, help the web", TechCrunch, April 12, available at: <https://techcrunch.com/2011/04/12/duolingo/> (accessed November 25, 2017).
- Siggelkow, N. (2007), "Persuasion with case studies", *Academy of Management Journal*, Vol. 50 No. 1, pp. 20-24.
- Sorescu, A. (2017), "Data-driven business model innovation", *Journal of Product Innovation Management*, Vol. 34 No. 5, pp. 691-696.

- Statista (2016), "Number of smartphone users worldwide from 2014 to 2020 (in billions)", available at: [www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/](http://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/) (accessed January 18, 2018).
- Stone, B. (2011), "Jawbone to release wristband health monitor", *Sfgate*, November 5, available at: [www.sfgate.com/business/article/Jawbone-to-release-wristband-health-monitor-2324160.php](http://www.sfgate.com/business/article/Jawbone-to-release-wristband-health-monitor-2324160.php) (accessed November 25, 2017).
- Täuscher, K. and Laudien, S.M. (2018), "Understanding platform business models: a mixed methods study of marketplaces", *European Management Journal*, Vol. 36 No. 3, pp. 319-329.
- Tempini, N. (2017), "Till data do us part: understanding data-based value creation in data-intensive infrastructures", *Information and Organization*, Vol. 27 No. 4, pp. 191-210.
- Trabucchi, D., Buganza, T. and Pellizzoni, E. (2017), "Give away your digital services: leveraging Big Data to capture value", *Research Technology Management*, Vol. 60 No. 2, pp. 43-52.
- Trabucchi, D., Buganza, T., Dell'Era, C. and Pellizzoni, E. (2018), "Exploring the inbound and outbound strategies enabled by user generated Big Data: evidence from leading smartphone applications", *Creativity and Innovation Management*, Vol. 27 No. 1, pp. 42-55.
- Troilo, G., De Luca, L.M. and Guenzi, P. (2017), "Linking data-rich environments with service innovation in incumbent firms: a conceptual framework and research propositions", *Journal of Product Innovation Management*, Vol. 34 No. 5, pp. 617-639.
- Vajjhala, N.R. and Ramollari, E. (2016), "Big Data using cloud computing-opportunities for small and medium-sized enterprises", *European Journal of Economics and Business Studies*, Vol. 4 No. 1, pp. 129-137.
- Von Ahn, L. (2013), "Duolingo now translating BuzzFeed and CNN", Duolingo Official Blog, available at: [www.duolingo.com/comment/954969/Duolingo-now-translating-BuzzFeed-and-CNN](http://www.duolingo.com/comment/954969/Duolingo-now-translating-BuzzFeed-and-CNN) (accessed November 25, 2017).
- Wamba-Fosso, S., Akter, S., Edwards, A., Chopin, G. and Gnanzou, D. (2015), "How 'Big Data' can make big impact: findings from a systematic review and a longitudinal case study", *International Journal of Production Economics*, Vol. 165, pp. 234-246.
- Xie, K., Wu, Y., Xiao, J. and Hu, Q. (2016), "Value co-creation between firms and customers: the role of Big Data-based cooperative assets", *Information & Management*, Vol. 53 No. 8, pp. 1034-1048.
- Yin, R.K. (2009), "How to do better case studies", *The SAGE Handbook of Applied Social Research Methods*, Vol. 2, Thousand Oaks, CA, pp. 254-282.

#### About the authors

Daniel Trabucchi is PhD Student at the School of Management of Politecnico di Milano, where he serves as Researcher of LEADIN'Lab, the Laboratory for Leadership, Design and Innovation. His research interests are focused in innovation management. In particular, he has been working on the impact of digital technologies and Big Data on two-sided market business model; moreover, he focuses on innovation strategy based on the interplay between technology and meaning. Daniel Trabucchi is the corresponding author and can be contacted at: [daniel.trabucchi@polimi.it](mailto:daniel.trabucchi@polimi.it)

Tommaso Buganza is Associate Professor of Leadership and Innovation at the School of Management of Politecnico di Milano where he also is Co-founder of LEADIN'Lab, the Laboratory for Leadership, Design and Innovation. He is responsible for the Project Management Academy and Coordinator of the Innovation and Training Area at MIP (Politecnico di Milano Graduate School of Business). He is Member of the Scientific Committee of the International Product Development Management Conference ELASM-IPDMC. His research activity explores the intersection between technological innovation and leadership and has been published in peer-reviewed journals.