Please cite this article as:

Trabucchi D, Buganza T, Dell'Era C and Pellizzoni E (2018).
Exploring the Inbound and Outbound Strategies enabled by User Generated Big Data: Evidences from Leading Smartphone Applications.
Creativity and Innovation Management, Vol. 27, Pp. 42-55.
(DOI: 10.1111/caim.12241)

Exploring the Inbound and Outbound Strategies enabled by User Generated Big Data: Evidences from Leading Smartphone Applications

[Special Issue: Big Data for Open Innovation – Unveiling Challenges and Opportunities]

Daniel TRABUCCHI (corresponding author)
Department of Management, Economics and Industrial Engineering,
Politecnico di Milano – Piazza L. da Vinci, 32 20133 Milano Italy
Tel: +39 02 2399 3967, Fax: +39 02 2399 2720
daniel.trabucchi@polimi.it

Tommaso BUGANZA
Department of Management, Economics and Industrial Engineering,
Politecnico di Milano – Piazza L. da Vinci, 32 20133 Milano Italy
Tel: +39 02 2399 2804, Fax: +39 02 2399 2720
tommaso.buganza@polimi.it

Claudio DELL'ERA
Department of Management, Economics and Industrial Engineering,
Politecnico di Milano – Piazza L. da Vinci, 32 20133 Milano Italy
Tel: +39 02 2399 2798, Fax: +39 02 2399 2720
claudio.dellera@polimi.it

Elena PELLIZZONI
Department of Management, Economics and Industrial Engineering,
Politecnico di Milano – Piazza L. da Vinci, 32 20133 Milano Italy
Tel: +39 02 2399 3967, Fax: +39 02 2399 2720
elena.pellizzoni@polimi.it

Exploring the Inbound and Outbound Strategies enabled by User Generated Big Data: Evidences from Leading Smartphone Applications

[Special Issue: Big Data for Open Innovation – Unveiling Challenges and Opportunities]

ABSTRACT

The Open Innovation paradigm has been increasingly considered as a relevant approach to innovation. Among the different sources, the end users are particularly meaningful. Scholars highlighted several methods and strategies to involve them in the innovation process by asking, observing and giving them the chance to actually co-create. Digital technologies are expanding the span of opportunities in this direction, gathering a huge amount and variety of data while the end user enjoy a digital product, these data can be named as User Generated Big Data (UGBD). The aim of this research is to understand whether UGBD can contribute in User Innovation and to highlight the enabled strategies to create value through them. Leveraging on a multiple case study (Twitter, Spotify, Strava and Deliveroo), the paper first classify UGBD among the methods to foster User Centered Innovation, second it defines two strategies to create value relying on UGBD. First, companies can leverage on a Using Data strategy – addressing both the end user or other player in the ecosystem - fostering service innovation through an inbound approach. Second, a Selling data strategy can be pursued, addressing new clients and fostering business model innovation, enlarging the company's value chain in an outbound perspective.

INTRODUCTION

Back in 2012, IBM (2012) estimated a daily data creation of 2.5 quintillion bytes, an unimaginable number. Other researches (IDC, 2014) foresee a 10x growth of the digital universe from 2013 to 2020, moving from 4.4 trillion of gigabytes to 44 trillion. Moreover, the same research foresees an increase in the useful data (if tagged and analyzed) from 22% to 37%, which opens several avenues for potential applications (IDC, 2014).

Data obviously represent a key source in each management process and collecting data is becoming much easier. When in 2001 Passur Aerospace, a business intelligence company

operating in the aviation industry, wanted to offer an ETA service collecting real time data about the flights, they had to install an expensive network of passive radars nearby airports (McAfee & Brynjolfsson, 2012). Nowadays, thanks to digital technologies, similar data retrieving activities are performed in a significant cheaper and more efficient way. In many cases data are even virtually free as they come as a side-effect of the consumption of existing products and services. For example, data about customers' previous orders, previous searches, shopping cart history and other products observed, are collected almost for free by Amazon and then used to make suggestions to customers or even to pre-ship the products to the nearest hub waiting for the order to be placed (anticipatory shipping) (Erevelles et al., 2016). This data driven value creation is commonly referred to as Big Data.

The term Big Data (BD) is very popular nowadays (Lohr, 2012, van der Meulen, 2016) and several definitions have been proposed (Manyika et al., 2011; Davenport, 2012; Fisher et al., 2012; Havens et al., 2012; Johnson, 2012). BD is much more than data analytics (Fosso Wamba et al., 2015), due to five key features: volume, velocity and variety (McAfee & Brynjolfsson, 2012), later expanded with value (due to the importance of extracting value from the available data) (Leventhal, 2013) and veracity (in order to underline the importance of quality data and the chance to trust them) (White, 2012).

BD can embed huge innovation opportunities that, in many cases, only need to be unveiled, both for researchers (e.g. George et al., 2016) and practitioners (e.g. Buganza et al., 2015). For example, Waze, the navigation app bought by Google in 2013 for US$ 1.3 billion, is a great example of this dynamic: its innovative service is based on maps constantly updated thanks to data provided by user community while they are using the navigation service (Buganza et al., 2015). This huge amount and variety of data can be considered one of the sources that companies can leverage in an Open Innovation paradigm (e.g Kaplan and Haenlein, 2010; Mount and Martinez, 2014; Cohen et al., 2016), considering sources such as Social Media during the entire innovation funnel (Mount and Martinez, 2014), or to implement open data strategies with several opportunities to foster innovation (Berrone et al., 2017).

OI is defined as a distributed innovation process that leverages the flow of knowledge crossing company boundaries (Chesbrough & Bogers, 2014). These inflows and outflows are respectively able to accelerate the internal development process and the external exploitation process (Chesbrough, 2003; Randhawa et al., 2016). OI enriches the traditional innovation funnel by removing a traditional barrier: ideas, technologies, and solutions from external environments are incorporated within the innovation process and the developed innovations can also be exploited outside the company (Chesbrough & Crowther, 2006; Enkel et al., 2009,

West & Bogers, 2014). Several sources of external knowledge are suggested by literature: users, suppliers and competitors can provide useful insights to the innovation process but also non-customers, non-suppliers, and partners from other industries can play a key role (Enkel et al., 2009).

Heavy attention has been dedicated by literature to users due to their relevance in innovation processes (Almirall & Wareham, 2008; Følstad, 2008; Raasch et al., 2008; Schuurman et al., 2011; Leminen et al., 2012). In this vein, it is interesting to observe that a considerable share of the existing Big Data is actually generated by users (e.g. Kaplan and Haenlein, 2010) interacting with existing digital products and services: User Generated Big Data (UGBD). UGBD can provide a deeper understanding about the customers, which suggests to consider them from a User-Centered innovation perspective (Kelley, 2001). In particular, opposite to the current techniques studied to collect data from users, e.g. interviews or user toolkits, digital technologies allow to retrieve huge amounts and variety of data. In this vein, UGBD can provide additional and more comprehensive insights from users and foster the innovation process.

Scholars already started investigating the role of UGBD through an open innovation perspective, for example studying how crowdsourcing can be leveraged as a tool for data analysis (Martinez and Walton, 2014) and how user generated data in social media can be considered a knowledge source for innovation (Kaplan and Haenlein, 2010; Mount and Martinez, 2014). Nevertheless, there is still the need to deep in this topic and to explore its potentialities (e.g. Martinez and Walton, 2014).

Considering Big Data generated by users as potential source of innovation, this paper aims at investigating what are the peculiarities that differentiate User Generated Big Data (UGBD) from other User Innovation techniques and how they can enable value creation strategies. More specifically, interpreting the UGBD as a potential evolution of the rich literature about User Innovation techniques, the paper aims at identifying its distinctive features in comparison with other techniques. Furthermore, the paper investigates the value creation strategies enabled by UGBD in order to provide additional insights about the opportunities supported by this emerging User Innovation technique.

Five sections follow this introduction. In section 2, the literature background provides theoretical grounds of the study. Section 3 describes the design of the research, presenting the empirical setting, the sample selection and the data gathering/analysis processes. Then empirical results are discussed in Section 4. In the discussion, Section 5, the answers to the

research questions are presented. Finally, Section 6 concludes and outlines avenues for future research.

LITERATURE BACKGROUND

As previously mentioned users can be interpreted as sources of innovation or, in a wider perspective, contributors to innovation. Open Innovation (OI) is defined as a distributed innovation process that leverages the flow of knowledge crossing company boundaries (Chesbrough & Bogers, 2014). Several sources of external knowledge are found in the literature stream about open innovation; users, suppliers and competitors represent just few of them. In the inbound open innovation, or outside-in process, the company enriches its own knowledge through external stimuli by monitoring the external environment (Laursen & Salter, 2006; Lettl et al., 2006; Piller & Walcher, 2006, West & Bogers, 2014). Outside there are tons of ideas, technologies, and solutions that can be useful in enriching internal innovations or solving internal innovation problems. In the outbound open innovation, or inside-out process, the company canalizes its innovations into external markets or organizations that better suit their diffusion and exploitation (Lichtenthaler & Ernst, 2007). The critical role that users can play in the innovation process has been underlined by several scholars (Almirall & Wareham, 2008; Følstad, 2008; Raasch et al., 2008; Schuurman et al., 2011; Leminen et al., 2012), to the point that we have seen a progressive evolution in this approach over the last two decades: from user-centered design (Kelley, 2001; Lojacono & Zaccai, 2004) to human-centered design (Buchanan, 2001), and from design thinking (Brown, 2008 and 2009; Martin, 2009) to design sprint (Knapp, 2016). While user-centered design emerged in the '90s as a referential paradigm based on a deep understanding of users' needs, human-centered design challenged the initial view in the late '90s proposing a broader view of the user: being immersed in her lifestyle and cultural context, the interpretation of the same product/service can significantly change across users; conceiving a new product/service around humans instead of users allow to properly consider emotional and symbolic values that otherwise can be underestimated (Buchanan, 2001; Kelley, 2001; Lojacono & Zaccai, 2004). Design thinking represents a formal method for practical, creative resolution of problems and creation of solutions, with the intent of a continuously improving results. Pioneers of design thinking Tim Brown and Roger Martin have spearheaded the shifting role of design in business from methods and approaches used only by designers to the method of creative action adapted for business purposes (Brown, 2009; Martin, 2009). A design sprint is

a five-phase framework that helps answer critical business questions through rapid prototyping and user testing (Knapp, 2016); conceived by Google Ventures, it merges design thinking with lean startup approach in order to accelerate the capability to deliver innovative solutions.

Focusing on a specific category of external knowledge source such as the users we explore two different literature streams that differently interpret the role users can play in supporting the development of innovations: User-Centered and User-Driven Innovation.

User-Centered Innovation

The main assumption of the user-centered innovation (UCI) approach is that a firm may infer unique insights to inform product innovation by asking users about their needs or, more effectively, by observing them as they use existing products, and by tracking their behavior in consumption processes. Users play a double role in the UCI paradigm: they are both the source of the information that a company has to know in order to innovate and the main beneficiaries to whom the final solutions are addressed (Leminen et al., 2012; Nyström et al., 2014). According to Karat (1996), UCI exploits data collected from users in order to generate new ideas and to assess the quality of the final solution. By providing valuable information, documenting their behaviors and testing the final solutions, users can significantly improve both the effectiveness and the efficiency of the innovation process. The evolution of the UCI approach has produced a parallel explosion in methods and techniques that can support its adoption.

Interviews represent the most traditional technique adopted in order to investigate users' needs (see Table 1). Although interviews are particularly efficient because they allow the collection of information about several users with few resources, this technique has several limitations. The user is aware of the ongoing investigation, even if she is only partially informed about the final objectives of the interview since only part of the information investigated by the interview can be disclosed. Questions may significantly influence the answers collected to the point that users can provide partial or fake answers to the investigation. Differently from other UCI techniques, interviews do not allow to take into consideration the influence that the context of use can generate on users' needs and do not consider the insights that might come from the interaction between different users. As underlined by Dahan and Hauser (2001) and Sanders (2002) this technique performs an explicit need in investigating and, as a consequence, mainly provides insights that can nurture incremental innovations. The *focus group* technique foresees the meeting of a representative

sample of users where participants are invited by the organizing company to discuss and share specific problems or issues (see Table 1). Obviously users are aware of the investigation process and, similarly to the interview technique, the role of context of use can be misleading: usually focus groups are organized in locations that do not replicate the real context of use or which eventually can be perceived by users as artificial. Differently from the interview technique, though, the contemporary participation of different users can facilitate the emergence of critical views; in fact, the possibility of leveraging alternative opinions can allow serendipitous solutions to be to identified. At the same time, as argued by Dahan and Hauser (2001), it could happen that social dynamics discourage participants from explicitly revealing their own needs in the presence of others; in other words, opinions expressed by leading participants can significantly influence the ideas of other participants. Finally, differently from the interview technique, focus groups usually rely on smaller numbers of users providing less robust results, even if they are usually richer. *Applied ethnography* is a UCI technique based on the observation of users in real-life settings with the aim of investigating needs demonstrated by behaviors and interactions (see Table 1). The main assumption behind this technique is that very often users are not completely aware of their behaviors in interactions with other people and products (Whitney & Kumar, 2003; Rosenthal & Capper, 2006). Unlike previous techniques, applied ethnography acts in real-life settings, taking into consideration the influence the context of use can have on users' behaviors and needs. Like the focus groups, applied ethnography also supports the contemporary observations of different users. This technique aims to discover the symptoms of unexpressed needs (Burns, 2000). One of the main limitations of applied ethnography is its cost and consequently the questionable robustness of the empirical results frequently associated with few observations; the support, however, provided by technologies, such as video recording and sensors, reduces significantly this limitation.

[Table 1]

User-Driven Innovation

According to Franke (2014), a new product can be interpreted as a user-driven innovation (UDI) if "it was invented and prototyped by an institution that aims to benefit from the innovation by using it, not by selling it". Users can be both individual consumers (e.g., the case of the snowboard investigated by Shah, 2000) or firms (e.g., the case of medical robot system for neurosurgery analyzed by Lettl et al., 2006). As underlined by Franke (2014), UDIs are an

ancient mode of innovation, but their relevance and diffusion significantly increased in the last decades for several reasons: the internet and social networks allow individuals to share their ideas and proposals much more easily than decades ago when geographical and social impediments significantly decreased the ability of people to move around the world and get in touch with other individuals. Open source software, such as Linux or Firefox, and digital platforms based on user-generated content, such as Wikipedia or YouTube, represent just few examples of UDIs. The UDI approach has gained increasing attention also from scholars. According to Franke (2014), in the period 1986–1990 only two papers appeared in peer-reviewed journals investigating UDI issues, while between 2006 and 2010 there were 60. UDI represents an emerging paradigm that is significantly transforming the way companies can innovate (Baldwin & von Hippel, 2011; Dahlander & Frederiksen, 2011; von Hippel et al., 2011).

According to Urban and von Hippel (1988), the *lead user* method is a managerial heuristic that enables companies to search for particularly attractive user innovations and identify radically new business opportunities (see Table 2). More specifically, this method requires researching, identifying and observing resourceful users who autonomously and spontaneously develop "ad hoc" solutions to better satisfy their personal needs. Several researches demonstrate that the lead user method supports the generation of ideas with attractive results from a commercial point of view (Urban & von Hippel, 1988; Herstatt & von Hippel, 1992; Olson & Bakke, 2001; Skiba et al., 2009). Von Hippel (von Hippel, 1998, 2001; von Hippel & Katz, 2002; de Jong & von Hippel, 2009) conceptualizes "*user toolkit[s]* as coordinated sets of design tools that allow individual users to self-design their own individual product according to their individual preferences and give visual and informational feedback on (virtual) interim solutions" (see Table 2). User toolkits allow manufacturers to abandon their attempts to understand user needs transferring instead the need-related aspects of products and services development directly to users through an appropriate toolkit (von Hippel, 2001). User toolkits are based on the idea that manufacturers possess the general knowledge of the solution possibilities, while the users possess the specific knowledge about needs. This information is sticky and therefore cannot be easily transferred from the user to the manufacturer. User toolkits can be used in a variety of settings and have been proved to be effective from the production of electronic circuitry to Apache security software (Franke & von Hippel, 2003). *Crowdsourcing* probably represents the most recent method that aims to profit from user creativity (see Table 2). As it is in its evolutionary stage, the concept of crowdsourcing can assume different forms: broadcast search (Jeppesen & Lakhani, 2010), innovation contest

(Terwiesch & Xu, 2008), virtual co-creation (Füller, 2010), innovation tournament (Terwiesch & Ulrich, 2009) or virtual customer environment (Nambisan, 2002). The method is based on an online question or challenge proposed by a company to the "crowd"; at the end of the call the company assesses the submitted proposals, selecting the best one and rewarding whoever submitted it (Nambisan, 2002; Ogawa & Piller, 2006; Dahlander & Magnusson, 2008; Terwiesch & Xu, 2008; Mustak et al., 2013; Saarijarvi et al., 2013; Pellizzoni et al, 2015). The crowdsourcing allows the collection of completely unexpected concepts considering that submitters can come from disparate knowledge domains and consequently can provide different perspectives on a problem (Jeppesen & Lakhani, 2010).

[Table 2]

As previously mentioned, digital technologies are supporting the production of an incredible amount of data in the world (i.e. big data). More specifically, exploiting the opportunities provided by digital technologies, companies are able to collect large amount and variety of data, directly generated by the users during the delivery of a service as a by-product (i.e. user generated big data). Even if the advantages and potentialities provided by UGBD are quite evident, the interpretation of them as an alternative user innovation technique can highlight the distinctive features they provide and consequently can guide managers in properly collect and exploit UGBD according to the challenge they are facing. For this reason, the paper aims at exploring the peculiarities of UGBD compared to other user innovation techniques in order to identify the peculiarities embedded in UGBD.

> **RQ1**: *What are the peculiarities that differentiate User Generated Big Data (UGBD) from other User Innovation techniques?*

Moreover, in order to clarify the managerial opportunities provided by UGBD the paper aims at identifying the strategies can be pursued in terms of value creation. More specifically, the paper aims at investigating how the collection of UGBD can support companies in pursuing different value creation strategies.

> **RQ2**: *What value creation strategies can be enabled by User Generated Big Data (UGBD)?*

RESEARCH DESIGN

A multiple case study method has been selected due to a lack of previous studies in the field. The adopted method is coherent with the exploratory intent of the research (Yin, 1984). By leveraging this research strategy, it is possible to develop a holistic and contextualized analysis in order to highlight the critical variables of a phenomenon (Eisenhardt & Graebner, 2007). We relied on secondary sources, using multiple case study with a retrospective intent (Yin, 1984).

Empirical Setting

The smartphone application industry represents a relevant empirical setting for the proposed research questions. The number of smartphone users is constantly growing; more than 2.87 billion people will own one of them by 2020 (Statista, 2016a). In June 2016, more than 2.2 million mobile apps were available for Android and 2 million for iOS, which represent the two largest app stores (Statista, 2016b). This will lead to more than 352 billion app downloads by 2021 (AppAnnie, 2017), with gross consumer spend in the different app store over $139 billion (AppAnnie, 2017). Smartphones are so much more than simple phones, as they have numerous sensors (e.g. compass, GPS tracker, accelerometer) embedded in them. Digital companies working in the field can then gather a huge amount and variety of data continuously. This empirical setting accomplishes the two main requirements to reach the research aim of this paper. First, smartphones are BD generators, and second, those data are user generated during the usage of a digital service, which is coherent with the kind of sources that can be implemented fostering UCI or UDI.

Case Studies Selection

The empirical research was conducted by following an inductive approach and using the case studies as inspiration for new ideas (Siggelkow, 2007). The case studies selection process aimed to find companies that have shown the ability to leverage on user generated data to foster innovation. More in detail, we selected the case studies according to two basic criteria: i) we searched for mobile apps that either leverage data actively provided by users (e.g. Social Media) or gather data leveraging the sensors embedded within the smartphone; ii) we selected apps with a high number of users in order to observe Big Data not only in terms of variety and velocity, but also in terms of pure volume (McAfee and Brynjolfsson, 2012). Through an iterative approach we analyzed seventeen mobile apps relying on secondary

sources, once we found different behaviors within our sample we searched for other applications that leveraged data in a similar way. The iterative process stopped when we stopped finding new ways to leverage user generated data to foster innovation in a grounded theory perspective (Corbin and Strauss, 2008). Finnnally, we built a convenient sample of four case studies aiming to show the greatest heterogeneity (for papers with similar kind of analysis see Dell'Era and Verganti, 2009; Abecassis-Moedas et al., 2012; Dell'Era et al., 2017). The main differences represented in the sample are in terms of *i)* app category (Social Network, Music, Health & Fitness and Food & Beverage), *ii)* phase of the lifecycle (start-up VS established companies) and *iii)* type of company (private VS public) in order to increase the replicability of the ways in which users can contribute to the innovation process by providing data and different ways to create value through them in different settings (see Table 3).

[TABLE 3]

Data Gathering and Data Analysis

The case studies rely mainly on secondary data leveraging on multiple sources, such as company official websites (e.g., general information on the service), app stores (e.g. Google Play for the download range) and several technology magazines (e.g., insights into the usage of the data, interviews), for a total of 33 documents and almost 300 pages (the breakdown of the data sources and of the documents on the single case is summarized in Table 4). All the data were updated on November 20th, 2016. Moreover, the general manager of one of the four cases (Deliveroo) has been interviewed twice. The two interviews lasted between 35 and 60 minutes and have been recorded and transcribed.

The analysis of the gathered documents and transcribed interviews has been done through an iterative process made of three main phases: reading, coding and interpreting (Saldaña, 2012). Following the recommendations of Corbin and Strauss (2008) we leveraged an *open coding process* (i.e. identifying key sentences from the documents and sorting them into first-order categories, as it happened - for example - with the quotes presented in the Empirical Results section). Then the categories have been combined through an *axial coding process* into higher-level categories by identifying the relationships between them and the analysed literature (e.g. key sentences from the gathered data have been aggregated in categories such as "Outbound open innovation strategy" or "Innovation addressing ecosystem partners"). Through this process, it has been possible to highlight i) how UGBD fits in the dimensions that

characterize UCI and UDI techniques and ii) which are the building elements of the strategies that these companies use to create value through UGBD.

The gathered data have been analyzed according to the dimensions defined through the literature review (which are summarized in Table 1&2), and through a comparison matrix of data as suggested by Miles and Huberman (1984).

[TABLE 4]

EMPIRICAL RESULTS

In the following sections, first the four cases are going to be presented in order to show how different companies leverage UGBD to create value.

Twitter – Tweets as enablers for Sentiment Analysis

Twitter is a social network founded by Jack Dorsey in 2006 in California. Users can post their tweets, 140 character updates, concerning what is going on in their lives. Users write about everyday problems, political ideas, and so on. Everything can be commented upon via this social network and using the right hashtag, linking together the 500 million tweets per day that the 310 million users write per day (Statista, 2016c); these messages are searchable all over the world. The community is the engine for the entire service: writing a tweet means creating value for the entire ecosystem, since other users will have the chance to re-tweet, answer and tweet again. The tag mechanisms, through hashtags, create threads related to specific topics.

The amount of data that Twitter gathers every second is incredibly huge; nevertheless, the data that the company owns is not just about tweets—which means text—but there is much more. Each single tweet is related to a specific moment in time, often to a geographical position, to a history of tweets (which can be a proxy of interests and opinions), and so on. Data gathered through the social network represents an incredible asset for the company itself which released their entire tweets' database for several years up to 2015 to third-party companies (i.e., Gnip, Datasift and NTT Data). Third-party companies, researchers and advertisers had the chance to use the user-generated data stream to analyze it and get insights and (a good proxy of) what the market says and thinks. In 2015, after the acquisition of Gnip, Twitter moved the data licensing within the company, thus removing any intermediaries. Zach Hofer-Shall, head of Twitter's ecosystem program, said, "*Direct relationships help Twitter*

develop an understanding of customer needs, get direct feedback for the product road map and work more closely with data customers to enable the best possible solutions for the brands that rely on Twitter data to make better decisions." (Hofer-Shall, 2015). The knowledge that can be created by leveraging these data can have different roles from understanding how the user's opinion is moving. Twitter Political Index proposed during the US presidential election a clear example of its application (Bilton, 2012; Patterson, 2016), and other examples are predictions about the stock market (Bollen et al., 2011), or the chance to target customers to reduce churns —as T-Mobile declared (van Rijmenam, 2013)—and many other potential applications. The previously mentioned UCI and UDI dimensions regarding Twitter are summarized in Table 5.

[Table 5]

Spotify – Play on-demand as enabler for insights into the market

Spotify was founded in 2006 by Daniel Ek and Martin Lorentzon in Stockholm, Sweden. The service was launched in 2008, announcing licensing deals with some of the many major music labels. Its revenue model is based on advertising in the free version and on monthly fees in the premium version, using a typical "freemium" model, and has more than 30 million paying subscribers (Hall, 2016). Besides giving access to millions of songs, Spotify is a community-based service; users can suggest songs to their friends, follow user-created playlists, and so on. All this leads to the creation of top-viral charts, along with traditional top-listened charts. Spotify is a data-driven company: they created Luigi, a Python framework for data flow definition and execution which they open-sourced, and is used to provide music suggestions and radio playlists from a user-centric perspective (Thelin, 2012). Being a data-driven company means having the chance to gather and mix together different sources of data (e.g., what users are listening to, when, where, when and what they skip). Spotify users create 600 gigabytes of data per day; they have more than 28 petabytes of storage in 4 data centers across the world. Interestingly, Spotify uses this huge amount of data not just to provide music suggestions. Spotify gives artists and managers access to data on how users listen to their music, in term of total streams, track-by-track information and demographic analyses (e.g., gender, location, and age); they have done this with Spotify Artists. "This is a new model, and it does take some time to get used to. By creating clarity around the model, and by discussing and explaining it, we hope we can overcome even some of the most vociferous critics. I hope this gets out some of the artists who won't engage with us. We can't be accused of hiding behind stuff: this is us being open and explaining our model". That was what Mark Williamson,

director of artist services at Spotify, told The Guardian (Dredge, 2013). An example of Spotify taking an innovative approach to BD is, "How students Listen" (Van Buskirk, 2014). They used aggregated data from Spotify student discount users to report how students listen to music in different schools, and found that NYU students like hipster music and University of Texas students like R&B and hate country music (Van Buskirk, 2014). Moreover, they have used streaming data to foresee the Grammy Awards winner since 2013, taking into consideration listeners' habits, album and track streaming, and so on, getting better rates than Billboard (Willens, 2015). These are just two simple examples, but behind them there is incredible value: Spotify owns data related to several aspects of music coming to fruition and they can be sold in the form of analytics to make music business decisions. The insights related to the UCI and UDI dimensions are summarized in Table 6.

[Table 6]

Strava – Rides as enablers for city planning

Strava is a fitness technology company founded in San Francisco in 2009, offering services through both a website and a mobile app. Strava is a community of athletes, amateurs or professionals, and through the app they can track activities, mainly cycling and running, but also other sports. Their users—estimated at 7 million (Everett, 2014)—let them gain more than 75 million rides yearly (Scott, 2015). Users are able to track their activities through the mobile app or from GPS devices (e.g., Garmin, Timex) and to get useful information, such as stats, average speed, and so on. They can compare their performances with past ones, but also with other athletes who did the same route, by leveraging community-based mechanisms. The social environment enables motivation and engagement mechanisms for the service and the fitness activity by following friends and other athletes.

They offer the basic service for free, while the premium service allows deeper analyses (goal setting, personal heat map, etc.), through a freemium business model. In just a few years Strava had a huge impact on sport, becoming a thriving social community merging the real and virtual world. The Oregon Department of Transportation (DOT) saw the incredible value hidden in this service: data. "*We were really deficient on the cycling and walking side of data*" said Margi Bradway (Davies, 2014), the transportation lead at the Oregon DOT, underlying how they created bike lanes where they looked logical, but without actually knowing which were the popular areas. In September 2013, Strava and the Oregon Department of Transportation closed a deal, giving birth to Strava Metro, a database coming from all the

runners and cyclists' activities, without personal information. They started selling data for $20,000 a year. Through this partnership the DOT can now understand where bikers speed up or slow down, where they stay on the street or go on the sidewalk, and so on. Strava created similar partnerships with others departments of transportation such as those in London, Glasgow, and others, thereby creating a new, unexpected revenue stream. More than 70 partners (cities and regions) are leveraging Strava data: according to Mark Shaw (Strava's technical chief officer), "*It helps show the return on investment, on the tax dollars being used by authorities for things like cycle lanes. They want to be able to show this was money well spent, or to learn that there was something they could have done better*". Creating heat maps was just a fun project, without seeing the value hidden within the data they were analyzing, and it turned out to be a business opportunity (Walker, 2016). The service was designed to help cyclists and runners get in touch with each other and share their own progress; now it can provide data to public administrations in order to answer specific questions, like how to make cycling safer. Table 7 summarizes the main dimensions coming from UCI and UDI reviews.

[Table 7]

Deliveroo – Orders as enablers for insights into operations and logistics

Deliveroo was founded by William Shu in London in 2013. It offers a delivery service that aims to change the final perception of this kind of service, triggered by the founder's direct experience when he moved to London from New York and was searching for delivery service from nice high-quality restaurants (Wood, 2016). The start-up is now working in 12 countries and more than 140 cities. After a certain number of investment rounds, it has recently been listed among the Unicorns, with a valuation higher that $1 billion, after a 650% growth in terms of deliveries in the last year, and with revenues growing monthly at a rate of 20-25% over the last three years (Dawson, 2016; Fedor, 2016; Olson, 2016; Wood, 2016).

Deliveroo can leverage different sources of data by working with different kinds of players. On the one hand, it gathers data from final customers, such as knowing what they buy, how often, from which restaurants, and so on. But this is just the beginning. On the other hand, Deliveroo also gathers data from restaurants, such as knowing the time needed from the moment they receive the order to the moment they send it out for delivery, as well as numerous statistics regarding the most popular dishes, seasonal information, and so on. In the end, the data gathered through the riders link the two sides: they know exactly how much time has been

used for the delivery and they can link this data with external sources, leveraging time and geo-localization. The company is able to create value by leveraging these data in different ways. First of all, a data scientist team works constantly on the algorithm that manages the links between final customers, riders and restaurants to increase the quality of the entire service. Shu declared: "If you tell the restaurant precisely when the [delivery] guy is going to be there, that's better than saying he'll be there in 10 minutes You give them precision. A Neapolitan pizza takes 90 seconds to cook. A steak takes 10 minutes. How do you pair that so the food is piping hot when the guy gets there? We have a million tests going on all over the place" (Olson, 2016). Moreover, they can leverage the gathered data to offer valuable information to the restaurants. One of the general managers of the company declared: *"[We offer them] suggestions on how to modify the dishes for the delivery, on their best dishes, on the composition of the packaging. […] We own a lot of interesting data for restaurants, and we periodically show them*." Moreover, he added, "*We can analyze the orders' trend of the restaurant and show them that seasonally they can see a decrease in their orders because a specific dish in no longer available. […] Another example can be working on the restaurant's metrics, such as preparation time, average waiting time, average receipt and give them different suggestions to improve all the logistics. Finally, there is the last topic, which is the most interesting and challenging, which means providing them suggestions on the areas where they should open a new restaurant, where we know there is a high potential request but a low offer coming from comparable restaurants*" [4]. Table 8 summarizes the main dimensions highlighted through UCI and UDI.

[Table 8]

DISCUSSION

As previously mentioned, this paper is aimed at understanding the peculiarities that differentiate UGBD from other User Innovation techniques and at investigating the value creation strategies enabled by UGBD. As a consequence, the discussion section is organized in accordance with the two research questions explored.

User Generated Big Data as an emerging User-Centered Innovation technique

In all case studies, users interact with the service and leave behind digital marks of their interaction. These marks can be different in nature (e.g., tweets, GPS tracks, music preferences, food preferences), but they share a set of common characteristics. First of all,

they are paid back by the service provider. More precisely they are needed to provide the service to the user. The mechanism that leads the user to provide those data is a typical *quid pro quo* one. Having my cycling stats would be not possible without allowing Strava to access my GPS data, additionally, having food delivered at home would be impossible without providing Deliveroo with info about the food I want and the place/time of the delivery. Customers have no problem leaving this digital information behind as it is necessary to enjoy the service. Moreover, this information has little value *per se* unless you join together large amounts of it. Companies in our sample proved to be capable of collecting and using this UGBD, coming from outside the company boundaries, to push innovation processes and to create new services for users and (in some cases) for different stakeholders, in a typical inbound OI dynamic.

[Table 9]

To provide an answer to our first research question (*RQ1: What are the peculiarities that differentiate User Generated Big Data (UGBD) from other User Innovation techniques?*) a comparative and more in depth analysis of the cases (see Table 9) is needed. As described in the literature section User-Centered Innovation and User-Driven Innovation techniques mainly differ on the contribution provided by the users: if UCI techniques aim at collecting insights about users' needs and behaviors, UDI techniques aim at engaging users in delivering solutions and concepts. In all the cases of the sample, rather than providing solutions, users provide insights: about their mood (Twitter), about their rides (Strava) or about their preferences (Spotify and Deliveroo). Thus, UGBD seem to be more likely a tool for User-Centered Innovation rather than a tool for User-Driven Innovation, since users have a double role as the sources of information and the main beneficiaries of the final solution (Leminen et al., 2012; Nyström et al., 2014). Nevertheless, the analyzed cases are clearly not examples of Interviews, Focus Groups or Applied Ethnography even if they share some common characteristics with all of them. Similar to interviews (Dahan and Hauser, 2001; Sanders, 2002), UGBD allow for the collection of huge quantities of data and for an increase in the generalizability of the results. Unlike interviews though, they allow to take into account the influence of the context on the contribution from the user and do not require a high level of awareness from the user (passive contribution). Similar to focus groups (Dahan and Hauser, 2001), UGBD allow for some interactions among users but, unlike them, they do not require an active role by the user, thus allowing for a higher generalizability of the results and for greater consideration of the influence of the context to be taken. Finally, like applied

ethnography (Whitney & Kumar, 2003; Rosenthal & Capper, 2006) they require a low awareness on the part of the user and take the context into account, but unlike applied ethnography, UGBD allow a much higher level of generalizability.

From a cross-case perspective we can see how two dimensions are constant: the level of awareness of users in contributing is low, while the context has an influence on the user's contribution. At the same time, the analyzed cases are different in terms kind of the insights received from the users (obliviously depending on the kind of service that is offered), but also in terms of interaction between users.

Indeed, if traditional UCI techniques can be classified through this dimension, this is not true for UGBD, as summarized in Table 10.

[Table 10]

As a result, we suggest UGBD as a potential new approach to User-Centered Innovation, similar to Applied Ethnography but more powerful to a certain extent, as its low cost and high replicability allow it to be spread it to a huge number of users. Obviously, the quantity and quality of data obtained through a physical observation are higher but it is important to note that current technologies already allow for a rich variety of data such as words, likes, GPS positioning, distance, speed, etc. to be gathered, and this variety is probably going to rapidly increase in the near future thanks to technological gadgets, like wearable sensors, etc.

This technique has some specific features: *i)* it leverages a variety of data directly created by the users (i.e. data tracked during the service or inserted by the user); *ii)* users are not completely aware of their contribution, decreasing the chance to bias the data (i.e. the data are gathered during the service itself in a non-invasive way); *iii)* the data are directly generated in the context where they are produced and *iv)* they may consider also the interaction among different users, due to the type of service.

These four features let UGBD emerge as a meaningful technique to study user needs and foster innovation through them. Indeed, it solves some of the issues that arise in the traditional techniques (e.g. the bias that the interviewer or the moderator in focus group may induce to the users, as well as the role of the researcher in applied ethnography that interpret qualitative data), increasing the generalizability through the high number of users and then of data. This technique may become more and more relevant, due to the pervasive diffusion of smartphones, and more in general digital technologies that allow companies to gather data in a continuous and non-pervasive way (Buganza et al., 2015).

Even if the UGBD can be interpreted as an emerging User-Centered Innovation because support the diffused collection of insights about users' needs and behaviors, it can also increase the effectiveness of User-Driven Innovation techniques. Supporting lead users with valuable information coming from UGBD or enriching the user toolkit with contextual information collected through UGBD modalities or providing crowdsourcing platforms with UGBD can significantly increase the potentialities of involved users in designing innovative concepts and delivering solutions.

Value creation strategies enabled by User Generated Big Data

The second research question (*RQ2: What value creation strategies can be enabled by User Generated Big Data (UGBD)?*) leads to two main strategic options adopted by companies in the sample: *Using Data* and *Selling Data* (see Table 11).

In the *Using Data* strategy is a clear example of inbound Open Innovation (e.g. Piller & Walcher, 2006, West & Bogers, 2014). UGBD are used to foster service innovation both enlarging the bundle of activities within the service (e.g., offering more services to customers in Deliveroo or listeners in Spotify) and moving to adjacent activity chains (e.g., restaurants in Deliveroo or artists in Spotify) (Sawhney et al., 2004).

Examples of UGBD *used* to enlarge the services offered to the final user can be the "most ordered dishes" in Deliveroo, which allows the end user to understand the most favored dishes of a restaurant, or the "customized suggested playlists" in Spotify (which gives the chance to do explore different and new music based on preferences of other users similar to you. On the other hand, examples of UGBD *used* to enlarge the services offered to other players in the ecosystem can be the insights that Deliveroo provides to restaurants on the customer satisfaction or on the average time they spend per order, or the insights that Spotify provides to artists on how their music is listened and (dis)liked.

In this perspective, the analyzed companies leverage an inbound strategy, meaning that they leverage knowledge and information created outside the boundaries of the company (i.e. through the users) to develop innovation within, through an out-side in process. This process may both enable innovation both towards the end-users or the ecosystem partners.

On one hand, UGBD can enable innovation on the existing service (as previously mentioned, enlarging the bundle of activities within the service), addressing the end-users. This means that the usage of UGBD has a direct impact on the people that generated them, receiving back an enhanced service or new features.

On the other hand, the use of UGBD can have a more complex impact of the company's mechanism. In particular, they might unveil and enable new opportunities to add value for some of the players involved in the ecosystems partners (as previously mentioned, moving to adjacent activity chains). Leveraging a multi-sided perspective (e.g. Parker et al., 2016; Trabucchi et al., 2017) UGBD can foster value creation not for those people who generated data, but for other partners directly involved within the main service (e.g. the artists or the music labels for Spotify and the restaurants for Deliveroo).

All cases in our sample apply a *Using Data* strategy to improve the services provided to the final users but in two cases (Twitter and Strava) a second *Selling Data* strategy allows them to find a lot more value in their data by serving very distant businesses. Twitter sells the content of public tweets for sentiment analysis and Strava sells anonymized data about bike rides and runs to a variety of city transportation departments around the globe.

The *Selling Data* strategy is a clear example of Outbound Open Innovation (e.g. Lichtenthaler & Ernst, 2007) and it seems to be an emergent one as in both cases it was not part of the initial value proposition. These two companies were able to see, after a while, a different value hidden within their data bases and to identify the potential customer who would made sense of that value. In this perspective, Twitter and Strava can be considered as platform providers (Gawer and Cusumano, 2014) that enable external parties to foster innovation starting from their data.

In this vein, the analyzed companies show how an outbound strategy related to UGBD is suitable as well. They work and rely on the knowledge (i.e. the data) they have within the company to foster innovation outside, unveiling new opportunities to other companies (e.g. companies doing market research in the case of Twitter, or municipalities and Departments of Transportation in the case of Strava), pursuing an inside-out strategy.

This strategy is probably replicable by every company that has got huge quantities of UGBD but it is obviously hard to implement and structure. So far, it seems more a matter of serendipity, that allows companies to innovate their business model moving to adjacent value chains. Business Model is a construct where the debate is still very open, see Massa and Tucci, 2017, for a clear and wide overview. We consider it through the definition proposed by Zott and Amit in 2010: "a system of interdependent activities that transcends the focal firm and spans its boundaries", p. 216. In this perspective, we consider the opportunity to broad the boundary of innovation (e.g. Casadesus-Masanell and Zhu, 2013; Massa and Tucci, 2014), without focusing on the creation of a new tangible product, but working as an information broker, curating and governing interactions and information flows (Choudary, 2015) among

different players within the ecosystem managed by the company, highlighting a new business opportunity creating value through UGBD and capturing it through a new revenues stream.
The two different strategies with their different impacts in terms of service innovation for end users or ecosystem partners and business model innovation addressing new clients are summarized in Table 11.

[Table 11]

CONCLUSIONS

This paper aims at investigating what are the peculiarities that differentiate User Generated Big Data (UGBD) from other User Innovation techniques and the related value creation strategies, addressing the need for research to deep the knowledge of Big Data within the Open Innovation paradigm as suggested by Martinez and Walton (2014).
The analysis of four different cases allows to identify two different strategies that companies can leverage to create value from UGBD, using both an inbound and an outbound OI approach (Chesbrough, 2003; Chesbrough & Crowther, 2006; West and Bogers, 2014; Randhawa et al., 2016). The first one is related to the chance to directly use the gathered data to increase the value delivered to the final user, or to other players within the product ecosystem. The second strategy allows companies to create value by enlarging their core business and selling data to external parties who can see in them an even greater value.
From a research perspective, this research classifies UGBD as a source for UCI (Leminen et al., 2012; Nyström et al., 2014), comparing it with the traditional techniques mentioned in the literature. This paper highlights UGBD value in terms of variety, purity of the gathered information (since users are not completely aware of their contribution), direct link with the context and the chance to consider interaction among different users.
From an open innovation perspective, UGBD can have a two-fold role. On one hand they can be used to implement an inbound strategy, enabling innovations that may target both the end users or the larger ecosystems of relation managed by the company. On the other hand, UGBD can be used to foster innovation through an outbound strategy, unveiling new opportunities enlarging to new clients. These findings build on previous researches that started showing the relevance of Big Data from an Open Innovation perspective (Kaplan and Haenlein, 2010; Mount and Martinez, 2014; Cohen et al., 2016)
From a practitioner stand-point, this research suggests three different ways to create value through UGBD that companies are already gathering through digital devices such as mobile

apps. In particular, UGBD may create value enabling different kinds of innovation addressing i) end users or ii) ecosystem partners with a service innovation perspective or iii) opening to new clients in a business model innovation perspective. These strategies may even become more and more relevant, due to the fast growing diffusion of smartphones, and more in general digital technologies, that allow companies to retrieve a huge amount and variety of data in a non-pervasive way.

This research doesn't mean to be exhaustive, but has an exploratory intent, indeed it is based on a small and convenient sample and leverages data gathered mainly through secondary sources. Moreover, the analysis focuses on a specific moment in time, without considering a longitudinal perspective. Furthermore, the research focuses on a single industry (mobile apps) and despite it considers different dimensions and typologies (through different categories), this lead to a low generalizability of the results.

The above mentioned limitations allow to highlight several avenues for further researches, for example enlarging the analysis to different industries, overcoming the mobile apps, to probe the findings in other contexts. Primary sources may be also involved to understand the antecedents of these strategies and to have a better view in terms of innovation process that created them.

Since the UGBD are getting more and more attention both from scholars and practitioners, future researches need to concentrate on the evolution of the phenomenon, that may unveil quiescent possibilities.

REFERENCES

Abecassis-Moedas, C., Ben Mahmoud-Jouini, S., Dell'Era, C., Manceau, D., & Verganti, R. (2012). Key Resources and Internationalization Modes of Creative Knowledge-Intensive Business Services: The Case of Design Consultancies. Creativity and innovation management, 21(3), 315-331.

Almirall E and Wareham J. (2008). Living labs and open innovation: Roles and applicability. The Electronic Journal for Virtual Organizations and Networks, Vol. 10, Pp. 21–26.

AppAnnie. (2017). App Annie Market Forecast 2016-2021. http://files.appannie.com.s3.amazonaws.com/reports/1704_Report_Market_Forecast_EN.pdf?mkt_tok=eyJpIjoiTVdaalpUaGtOakl5WmpreiIsInQiOiJMU3ZMRVdMOTU2VzAzNVRKN1U1WE1xUDhHbkp2S1YzODNtZjlEM2dPSndUVTZDcnRkRGIyYVpVdnd4NGRQcHF

Tc093eDBFNG9pSnRBeHlNYmVCZ1Y0TUh3XC85dlh5Z29lMjVRcFpYdWExXC9SUlhmRXgxVkhQT2RhZGFGcG80UWc5In0%3D [Accessed 8 April 2017]

Baldwin C and von Hippel E (2011). Modeling a Paradigm Shift: From Producer Innovation to User and Open Collaborative Innovation. Organization Science, Vol. 22, No. 6, Pp. 1399-1417.

Bilton, N. (2012) Twitter Unveils the Twindex, a New Political Index. The New York Times. http://bits.blogs.nytimes.com/2012/08/01/twitter-unveils-the-twindex-a-new-political-index/?_r=1 [Accessed 18 December 2016]

Bollen J, Mao H and Zeng X (2011). Twitter mood predicts the stock market. Journal of Computational Science, Vol. 2, No. 1, Pp. 1-8.

Brown T (2008). Design Thinking. Harvard Business Review, 84-92.

Brown T (2009). Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation. HarperCollins.

Buchanan R. (2001). Human dignity and human rights: Thoughts on the principles of human-centered design. Design issues, Vol. 17(3), Pp. 35-39.

Buganza, T., Dell'Era, C., Pellizzoni, E., Trabucchi, D., & 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, 24(3), 391-414.

Burns RB (2000). Introduction to research methods (4th ed.). Frenchs Forest, N.S.W.: Pearson Education.

Casadesus-Masanell, R., & Zhu, F. (2013). Business model innovation and competitive imitation: The case of sponsor-based business models. Strategic Manage- ment Journal, 34: 464–482.

Chesbrough, H. W. (2003) Open Innovation: The New Imperative for Creating and Profiting from Technology. Boston: Harvard Business School Publishing.

Chesbrough, H., and Bogers, M. (2014) Explicating Open Innovation: Clarifying an Emergin Paradigm for Understanding Innovation. In H. Chesbrough, W. Vanhaverbeke, & J. West, New Frontiers in Open Innovation (p. 4-28). Oxford: Oxford University Press.

Chesbrough, H., and Crowther, A. K. (2006) Beyond high tech: Early adopters of open innovation in other industries. R & D Management 36 (3), 229–236.

Choudary, S. P. (2015). Platform scale. Boston: Platform Thinking Labs.

Cohen, B., Almirall, E., & Chesbrough, H. (2016). The City as a Lab: Open Innovation Meets the Collaborative Economy. California Management Review, 59(1), 5-13.

Corbin, J. and Strauss. A. (2008) Basic of qualitative research: Techniques and procedures for developing grounded theory (3rd ed.). Thousand Oaks. CA. Sage Publications

Dahan E and Hauser JR (2001). Product development managing a dispersed process. In Handbook of Marketing. 2001.

Dahlander L and Frederiksen L (2011). The Core and Cosmopolitans: A Relational View of Innovation in User Communities. Organization Science, Vol. 23, No. 4, Pp. 988-1007.

Dahlander L and Magnusson M (2008). How Do Firms Make Use of Open Source Communities? Long Range Planning, Vol. 41, No. 6, Pp. 629-49.

Davenport, T. H. (2012). The Human Side of Big Data and High-Performance Analytics. International Institute for Analytics, 1–13.

Davies, A. (2014) Strava's Cycling App Is Helping Cities Build Better Bike Lanes. Wired http://www.wired.com/2014/06/strava-sells-cycling-data/ [Accessed 18 December 2016]

Dawson, F. (2016) Delivering 25% Month-On-Month Growth: The Deliveroo Story. Forbes. http://www.forbes.com/sites/freddiedawson/2016/02/26/delivering-25-month-on-month-growth-the-deliveroo-success-story/#2bb917554590 [Accessed 18 December 2016]

de Jong, J. P. J., and von Hippel, E. (2009) Transfers of user process innovations to process equipment producers: A study of Dutch high-tech firms. Research Policy, 38: 1181-1191.

Dell'Era, C., N. Altuna, S. Magistretti, and R. Verganti. (2017). Discovering quiescent meanings in technologies: exploring the design management practices that support the development of Technology Epiphanies. Technology Analysis and Strategic Management 29(2): 149-166.

Dell'Era, C., & Verganti, R. (2009). Design-driven laboratories: organization and strategy of laboratories specialized in the development of radical design-driven innovations. R&d Management, 39(1), 1-20.

Dredge, S. (2013) Spotify opens up analytics in effort to prove its worth to doubting musicians. The Guardian https://www.theguardian.com/technology/2013/dec/03/spotify-analytics-musicians-streaming-music-artists-earn [Accessed 18 December 2016]

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

Enkel, E., Gassmann, O. and Chesbrough, H. (2009) Open R&D and open innovation: Exploring the phenomenon. R & D Management 39 (4), 311–16.

Erevelles, S., Fukawa, N., and Swayne, L. (2016) Big Data consumer analytics and the transformation of marketing. Journal of Business Research, 69(2), 897-904.

Everett, D. (2014) How Strava continues to change cycling. Cyclingtips. http://cyclingtips.com.au/2014/12/how-strava-continues-to-change-cycling/ [Accessed 18 December 2016]

Fedor, L. (2016) Deliveroo technology provides opportunities beyond food. Financial Times. https://www.ft.com/content/d0ac352c-5a20-11e6-8d05-4eaa66292c32 [Accessed 18 December 2016]

Fisher, D., Drucker, S. M., and König, a. C. (2012) Exploratory visualization involving incremental, approximate database queries and uncertainty. IEEE Computer Graphics and Applications, 32(4), 55–62. http://doi.org/10.1109/MCG.2012.48

Følstad A. (2008). Living labs for innovation and development of communication technology: A literature review. The Electronic Journal for Virtual Organisations and Networks, Vol. 10, Pp. 99–131.

Fosso Wamba, 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, 165, 234–246.

Franke N (2014) User-driven Innovation. In (editors) Dodgson M, Gann DM and Philips N (2014). The Oxford Handbook of Innovation Management. Oxford University Press, Oxford.

Franke N and von Hippel E (2003). Satisfying heterogeneous user needs via innovation toolkits: The case of Apache security software. Research Policy, Vol. 32, Pp. 1199-1215.

Füller J (2010) Refining virtual co-creation from a consumer perspective. California Management Review, Vol. 52, No. 2, pp. 98–122.

Gawer, A., & Cusumano, M. A. (2014). Industry Platforms and Ecosystem Innovation. Journal of Product Innovation Management, 31(3), 417–433.

George, G., Osinga, E. C., Lavie, D., & Scott, B. A. (2016). Big data and data science methods for management research. Academy of Management Journal, 59(5), 1493-1507.

Hall, Z. (2016) Spotify now has 100M users, but only twice as many paid customers as Apple Music. 9t05Mac. https://9to5mac.com/2016/06/20/spotify-apple-music-users/ [Accessed 18 December 2016]

Havens, T.C., Bezdek, J.C., Leckie, C., Hall, L.O., and Palaniswami, M. (2012) Fuzzy c-means algorithms for very large data. Fuzzy Systems, IEEE Transactions on. 20(6). 1130–1146.

Herstatt C and von Hippel E (1992). From experience: Developing new product concepts via the lead user method: A case study in a "low-tech" field. Journal of Product Innovation Management, Vol. 9, Pp. 213-221.

Hofer-Shall, Z. (2015) Working Directly With the Twitter Data Ecosystem. GNIP. https://blog.gnip.com/twitter-data-ecosystem/ [Accessed 18 December 2016]

IBM. 2012.. What is Big Data? http://www-01.ibm.com/software/data/bigdata/ [accessed on 18 December 2016]

IDC. 2014. Data Growth, Business Opportunities, and the IT Imperatives. https://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm [accessed on 8 April 2017]

Jeppesen, L.B. and Lakhani, K.R. (2010) Marginality and problem solving effectiveness in broadcast research. Organization Science, 21(5), 1–18.

Johnson, J.E. (2012). Big Data + Big Analytics=Big Opportunity. Financial Executive. 28(6): 50-53.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. Business horizons, 53(1), 59-68.

Karat J. (1996). User-Centered Design: Quality or Quackery?. ACM/SIGCHI magazine, InteracGons july+august, Vol. 3, Issue 4, Pp. 18-20.

Kelley T (2001). The Art of Innovation. New York: Curreny.

Knapp J (2016). Sprint – How to solve big problems and test new ideas in just five days. Google Ventures.

Laursen, K., and Salter, A. (2006) Open for innovation: The role of openness in explaining innovation performance among U.K. manufacturing firms. Strategic Management Journal 27 (2), 131–50.

Leminen S, Westerlund M and Nyström AG. (2012). Living Labs as open innovation networks. Technology Innovation Management Review, Pp. 6–11.

Lettl, C., Herstatt, C., and Gemuenden, H. G. (2006) Users' contributions to radical innovation: evidence from four cases in the field of medical equipment technology. R&D Management, 36(3), 251-272.

Leventhal, R. (2013). Trend: Big Data analytics: from volume to value. Healthcare informatics: the business magazine for information and communication systems 30(2): 12-14.

Lichtenthaler, U., and Ernst, H. (2007) External technology commercialization in large firms: results of a quantitative benchmarking study. r&d Management, 37(5), 383-397.

Lohr, R. (2012) The Age of Bug Data. The New York Times. http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html [accessed on 18 December 2016]

Lojacono G and Zaccai G (2004) The evolution of the Design-Inspired Enterprise. Sloan Management Review, Spring, 75-79.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A.H. (2011). Big Data: the Next Frontier for Innovation, Competition and Productivity. McKinsey Global Institute.

Martin R (2009). The Design of Business: Why Design Thinking is the Next Competitive Advantage. Boston, MA: Harvard Business Press.

Martinez, M G, & Walton, B (2014). The wisdom of crowds: The potential of online communities as a tool for data analysis. Technovation, 34(4), 203-214.

Massa, L., & Tucci, C. L. (2014). Business model innovation. In M. Dodgson, D. M. Gann & N. Phillips (Eds.), The Oxford handbook of innovation management: 420–441. Oxford, UK: Oxford University Press.

Massa, L., Tucci, C. L. and Afuah, A. (2017) A critical assessment of business model research. Academy of Management Annals, 11(1), 73-104.

McAfee, A., and Brynjolfsson, E. (2012) Big data: the management revolution. Harvard Business Review, 90(10), 61–69.

Miles, M. B., and Huberman, A. M. (1984) Qualitative data analysis: An expanded sourcebook. London: Sage.

Mount, M., & Martinez, M. G. (2014). Social Media. California Management Review, 56(4), 124-143.

Mustak, M., Jaakkola, E. and Halinen, A. (2013), "Customer participation and value creation: a systematic review and research implications", Managing Service Quality, Vol. 23 No. 4, pp. 341-359.

Nambisan S (2002). Designing Virtual Customer Environments for New Product Development. Academy of Management Review, Vol. 27, No. 3, Pp. 392-413.

Nyström AG, Leminen S, Westerlund M and Kortelainen M. (2014). Actor role and role patterns influencing innovation in living labs. Industrial Marketing Management, Vol. 43, Pp. 483-495.

Ogawa S and Piller FT (2006). Reducing the risks of new product development. MIT Sloan Management Review, Vol. 47, No. 2, Pp. 65-71.

Olson EL and Bakke G (2001) Implementing the Lead User Method in a High Technology Firm: A Longitudinal Study of Intentions versus Actions. Journal of Product Innovation Management, Vol. 18, No. 2, Pp. 388-95.

Olson, P. (2016). Here's How Deliveroo Built An Army Of 5,000 Drivers In Just 3 Years. Forbes. http://www.forbes.com/sites/parmyolson/2016/02/17/deliveroo-army-5000-drivers-3-years/#712815497cec [Accessed 18 December 2016]

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.

Patterson, D. (2016) Clinton-Trump debate #1: What the data from social media tells us. TechRepublic http://www.techrepublic.com/article/clinton-trump-debate-1-what-data-from-social-media-tells-us/ [Accessed 18 December 2016]

Pellizzoni, E., Buganza, T., and Colombo, G. (2015) Motivation Orientations In Innovation Contests: Why People Participate. International Journal of Innovation Management, 19(04), 1550033.

Piller, F. T., and Walcher, D. (2006) Toolkits for idea competitions: A novel method to integrate users in new product development. R & D Management 36 (3), 307–18.

Raasch, C., Herstatt, C., & Lock, P. 2008. The dynamics of user innovation: Drivers and impediments of innovation activities. *International Journal of Innovation Management*, 12: 377-398.

Randhawa, K., Wilden, R., and Hohberger, J. (2016) A Bibliometric Review of Open Innovation: Setting a Research Agenda. Journal of Product Innoavation Management.

Rosenthal SR and Capper M (2006) Ethnographies in the Front End: Designing for enhanced Customer Experiences. Journal of Product Innovation Management, 23 (3), 215-237, May.

Sanders EB (2002) From user-centered to participatory design approaches. Design and the Social Sciences.

Saarijarvi, H., Kannan, P.K. and Kuusela, H. (2013), "Value co-creation: theoretical approaches and practical implications", *European Business Review*, Vol. 25 No. 1, pp. 6-19.

Saldaña, J. (2012) The coding manual for qualitative researchers (No. 14). Sage.

Sawhney, M., Balasubramanian, S., and Krishnan, V. V. (2004) Creating growth with services. MIT Sloan Management Review, 45(2), 34.

Schuurman D, De Moor K, De Marez L and Evens T. (2011) Living Lab research approach for mobile TV. Telematics and Informatics, Vol. 28, Pp. 271–282.

Scott, G. (2015) Strava 2014: the year in numbers. RoadCycling Guk. https://roadcyclinguk.com/sportive/strava-2014-year-numbers.html [Accessed 18 December 2016]

Shah S (2000) Sources and Patterns of Innovation in a Consumer Products Field: Innovations in Sporting Equipment. MIT Sloan School of Management. Working Paper No. 4105.

Skiba, F., & Herstatt, C. 2009. Users as sources for radical service innovations: Opportunities from collaboration with service lead users. *International Journal of Services Technology and Management*, 12: 317-337.

Siggelkow, N. (2007) Persuasion with case studies. Academy of management journal, 50(1), 20.

Statista. (2016a) Number of smartphone users worldwide from 2014 to 2020 (in billions). https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ [Accessed 18 December 2016]

Statista. (2016b) Number of apps available in leading app stores as of June 2016. https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/ [Accessed 18 December 2016]

Statista. (2016c) Number of monthly active Twitter users worldwide from 1st quarter 2010 to 3rd quarter 2016 (in millions). https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/ [Accessed 18 December 2016]

Terwiesch C and Ulrich KT (2009) Innovation Tournaments. Cambridge, MA: Harvard University Press.

Terwiesch C, Xu Y (2008) Innovation Contests, Open Innovation, and Multiagent Problem Solving. Management Science, Vol. 54, No. 9, Pp. 1529-43.

Thelin, M. (2012) Luigi is now open source: build complex pipelines of tasks. Spotify Developer. https://developer.spotify.com/news-stories/2012/09/24/hello-world/ [Accessed 18 December 2016]

Trabucchi, D., Buganza, T., & Pellizzoni, E. (2017). Give Away Your Digital Services: Leveraging Big Data to Capture Value. Research-Technology Management, 60(2), 43-52.

Urban GL and von Hippel E (1988). Lead user analyses for the development of new industrial products. Management Science, Vol. 34, Pp. 569-582.

Van Buskirk, E. (2014) Top 40 Musical Universities in America: How Students Listen. Spotify Insights http://static.echonest.com/HowStudentsListen/school.html?school=cornell%20university [Accessed 18 December 2016]

van der Meulen, R. (2016) Gartner Survey Reveals Investment in Big Data Is Up but Fewer Organizations Plan to Invest. http://www.gartner.com/newsroom/id/3466117 [accessed on 18 December 2016]

van Rijmenam, M. (2013) T-Mobile USA Cuts Downs Churn Rate By 50% With Big Data. DataFloq. https://datafloq.com/read/t-mobile-usa-cuts-downs-churn-rate-with-big-data/512 [Accessed 18 December 2016]

von Hippel E (1998) Economics of product development by users: The impact of `sticky' local information. Management Science, Vol. 44, No. 5, Pp. 629-44

von Hippel E (2001) PERSPECTIVE: User toolkits for innovation. Journal of Product Innovation Management, Vol. 18, 247-257.

von Hippel E and Katz R (2002). Shifting innovation to users via toolkits. Management Science, Vol. 48, Pp. 821-833.

von Hippel E, Ogawa S and De Jong JPJ (2011). The Age of the Consumer-Innovator. MIT Sloan Management Review, Vol. 53, No. 1, Pp. 27-35.

Walker, P. (2016). City planners tap into wealth of cycling data from Strava tracking app. The Guardian https://www.theguardian.com/lifeandstyle/2016/may/09/city-planners-cycling-data-strava-tracking-app [Accessed 18 December 2016]

West, J., and Bogers, M. (2014) Leveraging external sources of innovation: a review of research on open innovation. Journal of Product Innovation Management, 31(4), 814-831.

White, M. (2012) Digital workplaces Vision and reality. Business Information Review 29(4): 205-214.

Whitney P and Kumar V (2003) Faster, cheaper, deeper user research. Design Management Journal, Vol. 14, No. 2.

Willens, M. (2015) Shazam And Spotify's Grammy Picks Are More Accurate Than Billboard's. Here's Why. International Business Times. http://www.ibtimes.com/shazam-spotifys-grammy-picks-are-more-accurate-billboards-heres-why-1808170 [Accessed 18 December 2016]

Wood, Z. (2016) Deliveroo founder Will Shu drives company towards 'unicorn' status. The Guardian. https://www.theguardian.com/business/2016/aug/11/deliveroo-investment-company-unicorn-status [Accessed 18 December 2016]

Yin, R. K. (1984). Case Study Research, Design and Methods. London: Sage

Zott, C., & Amit, R. 2010. Business model design: An activity system perspective. Long Range Planning, 43: 216–226.

FOOTNOTES

[1] Through the Privacy Policy agreement (https://twitter.com/privacy?lang=en) the company declares the ways they can leverage on the gathered data

[2] Through the Privacy Policy agreement (https://www.spotify.com/it/legal/privacy-policy/?version=GB) the company declares the ways they can leverage on the gathered data

[3] Through the Privacy Policy agreement (https://www.strava.com/privacy) the company declares the ways they can leverage on the gathered data

[4] English translation of the original interview

[5] Through the Privacy Policy agreement (https://deliveroo.co.uk/privacy) the company declares the ways they can leverage on the gathered data

TABLES

Method/Technique	Contribution provided by users	Awareness of users in contributing	Influence of the context on users' contribution	Interaction between users
Interview	Insights about users' needs	High	Not controlled	No
Focus Group	Insights about users' needs	High	Artificial	Yes (few users)
Applied Ethnography	Insights about users' behaviors	Low	Real	Yes (few users)

Table1: Main techniques adopted in the User-Centered Innovation approach

Method/Technique	Contribution provided by users	Awareness of users in contributing	Influence of the context on users' contribution	Interaction between users
Lead User	Spontaneous solutions	Low	Real	No
User Toolkit	Supported solutions	High	Not controlled	No
Crowdsourcing	Potential concepts	High	Not controlled	No (usually)

Table 2: Main techniques adopted in the User-Driven Innovation approach

	Year and Country of foundation	Users (Millions)	Download (Millions)	App category	Phase of the lifecycle	Type of company
Twitter	2006 USA	310	500-1,000	Social network	Established company	Public
Spotify	2006 Sweden	100	100-500	Music	Established company	Private
Strava	2009 USA	8	5-10	Health and fitness	Start-up	Private
Deliveroo	2013 UK	4	0.5-1	Food delivery	Start-up	Private

Table 3: Case Studies Overview

	Official sources	Secondary sources	Number of documents
Twitter	Investor.Twitterinc.com Blog.Twitter.com	infoworld.com nytimes.com ibtimes.com brightplanet.com mashable.com statista.com	8
Spotify	Spotify.com Artist.Spotify.com Insights.Spotify.com	statista.com ibtimes.com Telegraph.co.uk 9to5mac.com	13

		theguardian.com promogogo.com	
 Strava	Strava.com Metro.strava.com Labs.Strava.com	theguardian.com cyclingtips.com wired.com	6
 Deliveroo	Deliveroo.com	fortune.com theguardian.com businessinsider.com techcrunch.com forbes.com	7 + 2 interviews

Table 4: Data sources and Data Gathering

	Contribution provided by users	**Awareness of users in contributing**	**Influence of the context on users' contribution**	**Interaction between users**
	Insights about Users' mood Opinions on various topics through tweets	Low They are aware of sharing an opinion within the community [1]	Real	Yes Through hashtags different threads are created, trending topics

Table 5: UCI and UDI dimensions in Twitter

	Contribution provided by users	Awareness of users in contributing	Influence of the context on users' contribution	Interaction between users
	Insights on Users' preferences Usage of the services (i.e. what they listen to, when, where, … -and suggestions to other users)	Low They are aware of contributing just suggesting song to other users [2]	Real	Yes Through suggestions and playlist sharing

Table 6: UCI and UDI dimensions in Spotify

	Contribution provided by users	Awareness of users in contributing	Influence of the context on users' contribution	Interaction between users
	Insights on Users' habits and statistics Usage of the services (i.e. rides and all related information)	Low They are not aware of contributing in something outside the usage of the service [3]	Real	Yes Through challenges

Table 7: UCI and UDI dimensions in Strava

	Contribution provided by users	Awareness of users in contributing	Influence of the context on users' contribution	Interaction between users
	Insights on Users' preferences Usage of the services (i.e. what they order, when they order it)	Low They are not aware of contributing in something outside the usage of the service [5]	Real	No Just indirectly through suggested dishes

Table 8: UCI and UDI dimensions in Deliveroo

	Contribution provided by users	Awareness of users in contributing	Influence of the context on users' contribution	Interaction between users

Twitter	Insights on Users' mood	Low	Real	Yes
Spotify	Insights on Users' preferences	Low	Real	Yes
Strava	Insights on Users' habits and statistics	Low	Real	Yes
Deliveroo	Insights on Users' preferences	Low	Real	No

Table 9: Comparison among Case Studies

	Contribution provided by users	**Awareness of users in contributing**	**Influence of the context on users' contribution**	**Interaction between users**
Interview	Insights about users' needs	High	Not controlled	No
Focus Group	Insights about users' needs	High	Artificial	Yes (few users)
Applied Ethnography	Insights about users' behaviors	Low	Real	Yes (few users)
User Generated Big Data	**Insights about Users' behaviors**	**Low**	**Real**	**Service dependent**

Table 10: UGBD as an emerging UCI technique (RQ1)

		USING DATA		SELLING DATA
Cases	**Gathered Data**	**Service** innovation addressing **End Users**	**Service** innovation addressing **Ecosystem Partners**	**Business Model** innovation addressing **New Clients**
Twitter	Tweets Hashtags Historical data Localization Time	Trending topics	-	Sentiment and market analyses
Spotify	Streams Suggestions Playlist Historical Data Localisation Time	Playlist	Insights to the artists	-
Strava	Starting point Breaks Journey Speed, Historical data Localization Time	Challenges	-	City plans and research
Delivero o	Orders Logistic measures Historical data Delivery time Historical data Localisation Time	Top ordered dishes	Insights to the restaurants	-

Table 11: Value creation strategies enables by UGBD (RQ2)