

FEDERAL UNIVERSITY OF RIO DE JANEIRO

COPPEAD GRADUATE SCHOOL OF BUSINESS

**RONY BARBOSA TOMAZ**

**BIG DATA ANALYTICS AS A SERVICE:  
How can services influence big data analytics capabilities in  
small and mid-sized companies?**

Rio de Janeiro

2020

**RONY BABOSA TOMAZ**

**BIG DATA ANALYTICS AS A SERVICE:  
How can services influence big data analytics capabilities in  
small and mid-sized companies?**

Master's dissertation presented to the COPPEAD Graduate School of Business, Universidade Federal do Rio de Janeiro, as part of the mandatory requirements in order to obtain the title of Master in Business Administration (M.Sc.).

Supervisor: Prof. Elaine Tavares, Ph.D.

Rio de Janeiro

2020

## CIP - Catalogação na Publicação

BB238b Barbosa Ttomaz, Rony  
BIG DATA ANALYTICS AS A SERVICE: How can services influence big data analytics capabilities in small and mid-sized companies? / Rony Barbosa Ttomaz. -- Rio de Janeiro, 2020.  
61 f.

Orientador: Elaine Tavares.  
Dissertação (mestrado) - Universidade Federal do Rio de Janeiro, Instituto COPPEAD de Administração, Programa de Pós-Graduação em Administração, 2020.

1. big-data-as-service. 2. big datat analytics capabilities. 3. big data. 4. big data analytics.  
I. Tavares, Elaine, orient. II. Título.

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a) autor(a), sob a responsabilidade de Miguel Romeu Amorim Neto - CRB-7/6283.

**RONY BARBOSA TOMAZ**

**BIG DATA ANALYTICS AS A SERVICE:**

**How can services influence big data analytics capabilities in  
small and mid-sized companies?**

Master's dissertation presented to the COPPEAD Graduate School of Business, Universidade Federal do Rio de Janeiro, as part of the mandatory requirements in order to obtain the title of Master in Business Administration (M.Sc.).

Approved on February 25, 2021

*Blaine NT Rodrigues*

\_\_\_\_\_  
**Elaine Maria Tavares Rodrigues, D.Sc - Advisor  
COPPEAD**

*Blaine NT Rodrigues*

\_\_\_\_\_  
**Marie Anne Macadar Moron, D.Sc  
COPPEAD**

*Blaine NT Rodrigues*

\_\_\_\_\_  
**Roberto da Costa Pimenta, D.Sc  
EBAPE/FGV**

Rio de Janeiro

2020

## ACKNOWLEDGMENTS

I want to acknowledge all who played a role in achieving this step of my life. First, I want to thank you, my parents, who always supported me in my educational and life journeys. Without the environment you created, encouraging education, and seeking my dreams, I would never get here.

Secondly, the COPPEAD members, the professors, all staff and my advisors, the Professors Elaine Tavares, the freedom and enthusiasm were fantastic, and the Professor Marie Anne Macadar. All the ideas, questions, and suggestions from both were invaluable.

Lastly, I want to express my special gratitude to John Kennedy, from 4property. Thank you for the support, the trust, and confidence deposited on me. I hope we can contribute to the companies better absorb technology with this study.

*"Every company has big data in its future, and every company will eventually be in the data business."*

Thomas H. Davenport

## RESUMO

Barbosa Tomaz, Rony. **Big Data Analytics as a Service: How can services influence big data analytics capabilities in small and mid-sized companies?** Rio de Janeiro, 2021. pp. 59  
Dissertation (Master's Degree in Business Administration) - COPPEAD Graduate School of Business, Federal University of Rio de Janeiro, Rio de Janeiro, 2021.

Big data analytics capabilities (BDAC) são um conjunto de rotinas facilmente reconhecidas, aprendidas, padronizadas e repetitivas que as organizações devem ter para executar e extrair valor de projetos de big data com sucesso. BDAC é comumente encontrado em grandes empresas que dispõem de recursos para desenvolver projetos de big data. As pequenas e médias empresas (PMEs), em contraste com as maiores, têm uma limitação clara de recursos para implementar projetos desta natureza. Este estudo explora como os provedores de serviços de análise de big data na nuvem podem impactar o BDAC em PMES. Ele foi realizado por meio do estudo de caso da plataforma iPPi, um recurso de Big Data Analytics, disponível na nuvem para o setor imobiliário na Irlanda. Foram explorados os pontos de vista do desenvolvedor da plataforma e de seus clientes. As descobertas sugerem impactos positivos nas capacidades técnicas e uma influência promissora na educação dos clientes do serviço sobre big data. O estudo mostra ainda que o desenvolvimento das BDAC vão além da mera aquisição do ferramental tecnológico, mas passa pelo desenvolvimento de uma cultura de orientação para dados nas PMEs.

Keywords: BDAC, SME, BIG DATA, BIG DATA ANALYTICS CAPABILITIES, DYNAMIC CAPABILITIES VIEW, BIG-DATA-AS-SERVICE

## ABSTRACT

Barbosa Tomaz, Rony. **Big Data Analytics as a Service: How can services influence big data analytics capabilities in small and mid-sized companies?** Rio de Janeiro, 2021. pp. 59  
Dissertation (Master's Degree in Business Administration) - COPPEAD Graduate School of Business, Federal University of Rio de Janeiro, Rio de Janeiro, 2021.

Big data analytics capabilities (BDAC) are a set of routines that are easily recognized, learned, patterned, and repetitious that organizations must have to execute and extract value from big data projects successfully. BDAC is mostly present in big companies that have the resources to develop big data projects. In contrast with big enterprises, SME's has an explicit limitation of resources to implement projects of this nature. This study explores how providers of big data analytics services in the cloud can impact BDAC on SMEs. Through the iPPi platform's case study, a study was carried about a Big Data Analytics tool available in the cloud for Ireland's real estate sector. The views of the platform developer and its customers were explored. The findings suggest positive impacts on technical capabilities and a good influence on educating service customers about big data. The study also shows that the development of BDACs goes beyond the mere acquisition of technological tools but includes developing a culture of data orientation in SMEs.

Keywords: BDAC, SME, BIG DATA, BIG DATA ANALYTICS CAPABILITIES, DYNAMIC CAPABILITIES VIEW, BIG-DATA-AS-SERVICE



## LIST OF ILLUSTRATIONS

Figure 1 - Characteristics of RBV, MBV, and Dynamic Capabilities .....	21
Figure 2 - Sample of aggregated data of a property .....	36
Figure 3 - Platform dashboard with a compilation of the data available.....	37
Figure 4 - Tool for property valuations .....	38
Figure 5 - Advertisement of the data analytics platform .....	39

## LIST OF TABLES

Table 1 - Definitions of big data analytics .....	17
Table 2 - BDAC typology overview.....	23
Table 3 - Data collection protocol .....	28
Table 4 - Companies represented on the Interviews.....	30
Table 5 - General overview of the company's participants of this study .....	33

## LIST OF ABBREVIATIONS

API	Application Programming Interface
BDA	Bid Data Analytics
BDAC	Big Data Analytics Capabilities
CEO	Chief Executive Officer
CTO	Chief Technology Officer
CRM	Customer Relationship Management
DCV	Dynamic Capabilities View
GDP	Gross Domestic Product
IT	Information Technology
NoSQL	Not Only SQL
OECD	Organization for Economic Co-operation and Development
SMEs	Small and medium-sized enterprises
SDK	Software Development Kit

## CONTENTS

<b>1. INTRODUCTION .....</b>	<b>13</b>
<b>1.1 ORIGIN AND RELEVANCE OF THIS STUDY .....</b>	<b>13</b>
<b>1.2 RESEARCH QUESTION.....</b>	<b>15</b>
<b>1.3 OBJECTIVES .....</b>	<b>15</b>
<b>2. LITERATURE REVIEW .....</b>	<b>16</b>
<b>2.1 BIG DATA .....</b>	<b>16</b>
<b>2.2 BIG DATA ANALYTICS (BDA) .....</b>	<b>17</b>
<b>2.3 BIG DATA AS A SERVICE .....</b>	<b>17</b>
<b>2.4 DYNAMIC CAPABILITIES VIEW (DCV) .....</b>	<b>18</b>
<b>2.5 BIG DATA ANALYTICS CAPABILITIES (BDAC).....</b>	<b>22</b>
<b>3. METHODOLOGY .....</b>	<b>25</b>
<b>3.1 THE CASE STUDY .....</b>	<b>25</b>
<b>3.2 DATA COLLECTION AND ANALYSIS.....</b>	<b>27</b>
<b>3.3 OVERVIEW OF THE SUBJECTS .....</b>	<b>30</b>
<b>Company C – International real estate agency - franchise.....</b>	<b>31</b>
<b>Company F – Local real estate agency .....</b>	<b>32</b>
<b>Company G – Local real estate agency.....</b>	<b>33</b>
<b>Company H – Local real estate consultancy company.....</b>	<b>33</b>
<b>3.4 GENERAL OVERVIEW OF THE COMPANIES .....</b>	<b>33</b>
<b>4. ANALYSIS AND DISCUSSION.....</b>	<b>35</b>
<b>4.1 OVERVIEW OF THE SERVICE PROVIDER .....</b>	<b>35</b>
<b>4.2 ANALYSIS OF THE DATA CAPABILITIES .....</b>	<b>39</b>
<b>4.3 ANALYSIS OF THE TECHNOLOGY CAPABILITIES.....</b>	<b>41</b>
<b>4.4 ANALYSIS OF THE PEOPLE CAPABILITIES .....</b>	<b>42</b>
<b>4.5 ANALYSIS OF THE PROCESS CAPABILITIES.....</b>	<b>45</b>
<b>4.6 ANALYSIS OF THE ORGANIZATION CAPABILITIES.....</b>	<b>46</b>
<b>5. CONCLUSION .....</b>	<b>48</b>
<b>5.1 LIMITATIONS.....</b>	<b>50</b>

<b>5.2 FUTURE STUDIES.....</b>	<b>50</b>
<b>REFERENCES .....</b>	<b>52</b>
<b>APPENDIX .....</b>	<b>58</b>
<b>Appendix A – Interview Script Customers .....</b>	<b>58</b>
<b>Appendix B – Interview Script Service Provider.....</b>	<b>60</b>

## **1. INTRODUCTION**

### **1.1 ORIGIN AND RELEVANCE OF THIS STUDY**

Knowledge, the fast speed of digitalization, information dissemination, and competition have transformed businesses recently (Bouwman et al., 2018). Data storage costs have reduced dramatically, and access to technology solutions became more democratic than before. People are continually generating much data through their connected devices (EMC, 2014).

The interaction between this vast amount of generated information and the capacity to extract value from this data is related to the Big Data concept. With the development of technologies that capture and store a tremendous amount of data, techniques and data analysis tools have also evolved, using data from strategic and operational perspectives. Organizations can improve their services by analyzing customer's purchases; enhance market strategies through marketing campaign analysis; use a more evidence-based knowledge co-creation of products, and achieving sustainability through long-term risk analysis (Wang, Shouhong & Wang, Hai, 2020).

The benefits of Big Data are still more readily available for larger firms. They concentrate the market share, customers, and profits away from those with no data (McAfee, 2012). Even though there is some availability access to data in a more democratic way, these companies still have some monopoly of restricted data, analytical skills, confidence, and internal resources (O'Connor, C. and Kelly, S., 2017).

In this scenario, small and medium-sized enterprises (SMEs) are inserted. These companies contribute over 50 percent to gross domestic product (GDP) and around two-thirds to formal employment (World SME Forum, 2017). They act as the main economic growth element by creating job opportunities, being innovative and productive (Maroufkhani et al., 2020; OECD, 2017; Singh, 2019; Talebi et al., 2012). In a globalized environment, one of the effective methods to reduce poverty and inequality in developing countries is developing and improving SMEs' position in the marketplace (Maroufkhani et al., 2020; Talebi et al., 2012). SMEs are an essential pillar of the economy.

Given the importance of SMEs, the adoption of Big Data and Big Data Analytics (BDA) is a critical topic. The development of analytical capabilities is crucial for these enterprises. Competition is one of the main reasons SMEs must adopt BDA (Tien et al.,

2020). The use of BDA can enable SMEs to have better performance, as they are more flexible than large companies (Ogbuokiri et al., 2015). The recent events during the COVID-19 lockdown also brought up the importance of using more big data analytics in post-pandemic decision-making. BDA technologies can help companies dealing with the kind of phenomena that companies may struggle with during lockdown: rapidly shifting consumer demand and confidence, operational disruptions, uncertainty, and redundancies in the workforce (OECD, 2020).

SMEs lagged in using BDA technologies (Coleman et al., 2016; Tien et al., 2020). These companies are often limited by scarce financial and organizational resources (Vajjhala & Ramollari, 2016) and a lack of understanding of how businesses deal with the process of BDA implementation and value generation (Mikalef et al., 2019).

Despite this, SMEs can access DBA technologies through cloud computing solutions. Cloud computing technologies enable a business to rent computing infrastructure and software services from a third-party provider without upfront investment in ICT capital (Marston et al., 2011; OECD, 2017b; OECD 2014; OECD 2019; Venters and Whitley, 2012). In the context of Big Data Analytics, there are cloud solution providers of big data analytics solutions. These companies own the expertise and manage the infrastructure, offering the results and tools to access information and insights more affordably. However, big data and big data analytics are complex topics. Even having some access to the technology, the company needs other internal capabilities to extract value from the big data. For example, companies must have data science skills or understand and incentivize managers to foster this culture.

In this context, this research proposes a case study on a cloud solution provider, iPPi, that offers big data analytics as a service for small and mid-sized real estate companies in Ireland. Most of its clients had no expertise built on big data analytics before using the platform. The service provider's goal is to bundle most of the technical aspects needed for big data analytics in an environment where the available data is abundant. They provide data and analytical capabilities for their clients via a cloud platform.

The analysis of this platform that provides big data analytics as a service was done through the Dynamic Capabilities View (DCV) lens, also using the foundations of the Big Data Analytics Capabilities (BDAC) framework developed by Mikalef et al. (2017) as a guide.

## **1.2 RESEARCH QUESTION**

This study's research question is: How big data analytics platforms influence and build big data analytics capabilities at SMEs?

## **1.3 OBJECTIVES**

This study's main objective is to understand how a big data analytics platform can influence SMEs' big data analytics capabilities. It also aims to understand the impact of these companies' newly acquired capabilities and their challenges.

It contributes to the theme covering the initial phases of building such analytical capabilities on companies and exploring the platform lens's theme instead of BDA capabilities developed only by internal resources.

It also has the following intermediary goals:

- Grasp on how companies can extract value from big data and how this development may occur;
- Explore the viability of data democratization through services; and
- Identify the expectations of SME's company owners/managers with big data.



## 2. LITERATURE REVIEW

This literature review aims to search for existing literature that approaches Big Data and business theories surrounding this subject. We explored several types of research on Big Data Analytics Capabilities and how these capabilities impact business performance. Basic concepts as the definitions of Big Data, Big Data Analytics, and Big Data Analytics as a Service were also reviewed. The compilation of the ideas is presented below.

### 2.1 BIG DATA

To start the study, we need initially to understand what Big Data is. The literature offers several definitions for Big Data. Mikalef et al. (2017), in their literature review, bring up 14 definitions of the term big data. Several explanations were analyzed to guide this study. In sum, Big data can be defined by the following characteristics, the called 7V's, that was built from an incremental contribution of the different authors along the time:

**Volume** - Represents the sheer size of the dataset due to the aggregation of many variables and an even larger set of observations for each variable (George et al., 2016);

**Velocity** - Reflects the speed at which data are collected and analyzed, whether in real-time or near real-time from sensors, sales, transactions, social media posts, and sentiment data for breaking news and social trends (George et al., 2016);

**Veracity** – Veracity ensures that the data used are trusted, authentic, and protected from unauthorized access and modification (Demchenko et al., 2013);

**Variety** – In big data, veracity comes from the plurality of structured and unstructured data sources such as text, videos, networks, and graphics, among others (George et al., 2016);

**Value** - Value represents the extent to which big data generates economically worthy insights and benefits through extraction and transformation (Wamba et al., 2015);

**Variability** - Concerns how insight from media constantly changes as the same information is interpreted differently, or news feeds from other sources help to shape a different outcome (Seddon and Currie, 2017);

**Visualization** - It can be described as interpreting the patterns and trends present in the data (Seddon and Currie, 2017).

## 2.2 BIG DATA ANALYTICS (BDA)

There are different ways to define big data analytics. Some authors do it by its characteristics; others also extend the definition by characteristics and include the analytical procedures, tools, and techniques that are employed; some authors even add up the impact that the analysis and presentation of big data can yield in terms of business value (Mikalef et al., 2017). We presented below some definitions of big data analytics:

**Table 1 - Definitions of big data analytics**

<b>Authors</b>	<b>Definition</b>
Loebbecke and Picot (2015)	Big data analytics: a means to analyze and interpret any kind of digital information. Technical and analytical advancements in BDA, which—in large part—determine the functional scope of today’s digital products and services, are crucial for developing sophisticated artificial intelligence, cognitive computing capabilities, and business intelligence.
Kwon et al. (2014)	Big data analytics: technologies (e.g., database and data mining tools) and techniques (e.g., analytical methods) that a company can employ to analyze large-scale, complex data for various applications intended to augment firm performance in multiple dimensions.
Ghasemaghaei et al. (2015)	Big data analytics, defined as tools and processes often applied to large and disperse datasets for obtaining meaningful insights, has received much attention in IS research <u>given its capacity to improve organizational performance</u>
Lamba and Dubey (2015)	Big data analytics is defined as applying multiple analytic methods that address the diversity of big data to provide actionable descriptive, predictive, and prescriptive results.
Müller et al. (2016)	Big data analytics: the statistical modeling of large, diverse, and dynamic datasets of user-generated content and digital traces.

Source: Mikalef et al., 2017

The definition from Lamba and Dubey (2015) of big data analytics was chosen as the anchor for this study. This definition is simple and comprehensive. The authors define Big Data Analytics as “the application of multiple analytic methods that address the diversity of big data to provide actionable descriptive, predictive, and prescriptive results”.

## 2.3 BIG DATA AS A SERVICE

The service model in technology is common sense. With the evolution of the internet speeds, most of the physical aspects of information technology moved to cloud computing and offered as services. Cloud computing can be succinctly defined as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., network servers, storage, applications, and services) that can be

rapidly provisioned and released with minimal management effort or service provider interaction (Mell, P., & Grance, T., 2011).

Cloud computing services are fundamental for big data projects. Big data involves a high volume of information in increased velocity and or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery, and process optimization (Beyer, M. A., Laney, D., 2012). To extract value from big data in an efficient way is a problem that requires complex structures. Moreover, this is especially true for a field that is in constant evolution, and new tools, solutions, and techniques arise on a day-to-day basis. To fulfill the potential of service-orientation is necessary to develop unique technologies to process large quantities of data within acceptable processing time effectively and easy access to the analysis (Zheng, Z., Zhu, J., & Lyu, M. R., 2013). The big data-as-a-service encapsulate all this in a bundle of a data storage solution (usually cloud), management, and analytics techniques into services and provides common big data-related services to users via programmable APIs and or visual interfaces, which greatly enhances efficiency, reduce costs, and enables seamless integration (Zheng et al., 2013).

Zheng et al. (2013) created a conceptual framework of Big Data-as-a-Service, which involves three layers: Big Data Infrastructure-as-a-Service, Big Data Platform-as-a-Service, and Big Data Analytics Software-as-a-Service. The integration between these three layers of services results in the value creation of the Big Data-as-a-Service. Big Data Infrastructure-as-a-Service is the layer that provides necessary computing and storage capacity for big data. The Big Data Platform-as-a-Service allows users to access, analyze and build analytic applications on top of large data sets (Horey et al., 2012). Furthermore, the Big Data Analytics Software-as-a-Service is the process of examining large amounts of data of various types to uncover hidden patterns, unknown correlations, and other useful information (Rouse, M., 2012).

## **2.4 DYNAMIC CAPABILITIES VIEW (DCV)**

This study is based on a Dynamic Capabilities View business strategy theory, an idea emerging from the most recent publications on the business area.

The different schools of strategy can be organized into two big classes. The first one is the early strategic theories that focused mainly on the environment outside of the company as the elements that should guide the company strategy. The second one, which emerged decades later, had a look shifted to a resource-based view, where the important factors were inside the

company (Carvalho, 2019). Below is a brief review of each theory and its main characteristics to provide a better context until we finally reach out to the Dynamic Capabilities view definition.

The early theories around business strategy views argued that external market factors shape a firm's performance. J.S Bain (Bain, J. S., 1968). developed the Structure-Conduct-Performance (SCP) paradigm, which stated that four elements impact the market structure: Supply (number of competitors), demand, product differentiation, and entry barriers. For Bain, the company's performance is a consequence of its environment (Carvalho, 2019). Later, Porter's work "How Competitive Forces Shape Strategy" introduced the five force models: the threat of new entrants, the bargaining power of customers, the threat of substitute products and services, the bargaining power of suppliers, and the competition for a position among competitors in its industry (Porter, 1979). Porter is another example of this class of theories that focus on external factors to explain firms' success.

The second kind of theories was formally introduced for the business literature by Birger Wernerfelt in his article "A Resource-based View of the Firm" in 1984, even though there was already literature focusing on companies' resources as factors that contribute to their growth, as the famous "The Theory of the Growth", 1959, wrote by Edith Penrose. A resource is "tangible and intangible assets tied semi-permanently to the firm" defined by Wernerfelt. The goal of Wernerfelt's paper is to identify "which circumstances will a resource lead to high returns over longer periods" (Wernerfelt, 1984, p. 172). Another important author of the resource-based strategic management theory is Jay Barney; in his article "Firm Resources and Sustained Competitive Advantage", 1991, brings up the investigation of conditions that help resources yield high returns and lead these resources to become a source of a sustained competitive advantage (Barney, 1991). It is important to cite Barney since he expands the definition of resource to "all assets, capabilities, organizational processes, firms' attributes, information, knowledge", and that for these resources to be a potential source of sustained competitive advantage, they must be valuable, rare, imperfectly imitable and must not have an equivalent substitute (Barney, 1991).

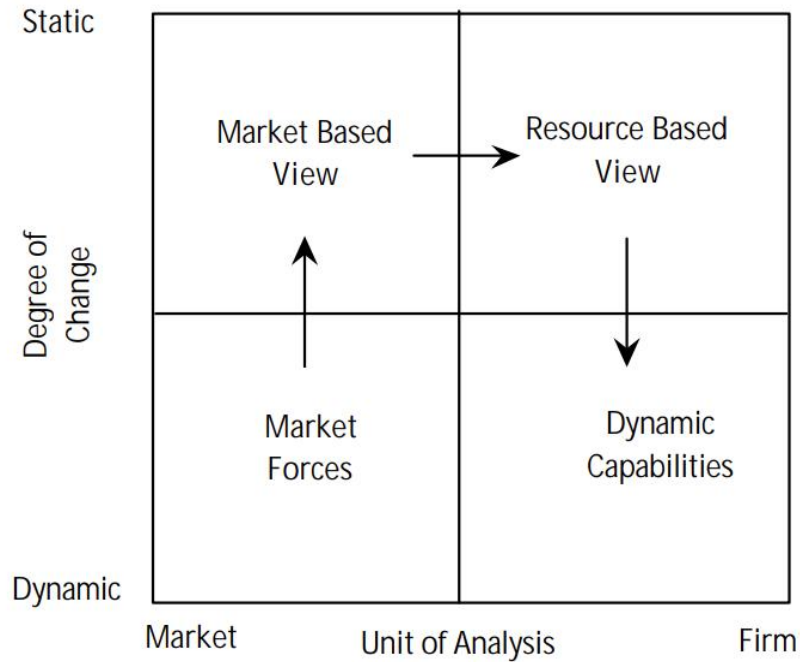
There was also a specific branch of the Resource-based view, called Knowledge-based View, as its first author also suggests, Robert M. Grant (Grant, 1996), it is an outgrowth of the resource-based approach. In his paper, he explores the various mechanisms by which firms coordinate this specialist knowledge that its members possess. They classify knowledge in two forms, explicit knowledge, or "Knowing about facts and theories", and tacit knowledge,

the “knowing how”. He exposes that explicit knowledge is much more transferable than tacit knowledge, which is much more costly and time-consuming to transfer (Grant, 1996, p. 111). He emphasizes the importance of the scope of knowledge to a capability and how it contributes to the resource-based view's inimitability factor. As broader and more integrated with the knowledge within a capability, the more difficult an imitation becomes. This situation is a source of casual ambiguity and replication barriers (Grant, 1996, p 117). Grant also touches on an interesting point that can be explored further for this proposed study. He brings up that knowledge is often built from strategic alliances. Companies need to reach their boundaries for knowledge and innovation (What brings us the justification of the relationship between technological service providers and clients that do not possess such expertise and knowledge).

The Dynamic Capability view is part of the theoretical perspectives that use the firm as the analysis unit. The approach was first presented by Teece, Pisano, and Shuen (1997), using the resource-based view fundamentals, but with the difference that the dynamic capabilities stress “exploiting internal and external, firm-specific competences to address changing environments”. According to Eisenhardt and Martin (2000), most of the authors describe DCV as a set of identifiable and specific routines that have often been the subject of extensive empirical research. Such routines are easily recognized, learned, patterned, and repetitious, directed towards independent corporate actions (Winter, 2003). These sets of routines are often defined as capabilities. However, it is important to stress that these routines are not ad-hoc events. Nevertheless, they are in constant adaptation and changes (Henderson and Venkatraman, 1993).

To have a better understanding of the difference between the different schools of thought, Pankaj M. Madhani (Madhani, 2010) presents the characteristics of the other theories based on two dimensions, the degree of change of the market and which unit of analysis the theory focus, the firm, or the market. He names the first school of thought presented above as Market Based view. The result of these differences is shown in the figure below.

**Figure 1 - Characteristics of RBV, MBV, and Dynamic Capabilities**



Source: Madhani, 2010

The reason to choose DCV as the foundation of the study is based on the fact that big data is a theme developed in a context where technology is evolving fast, and companies are dynamically evolving their process and culture to new trends. Markets are changing with the impact of technology. The unit of analysis is the firm. The theoretical groundings of RBV are also going to be used since it provides solid foundations for all relevant resources and how to evaluate their importance, while DCV enables us to examine the organizational capabilities and how they should be directed to achieve competitive performance gains by using big data in this environment of constant changes (Mikalef et al., 2016)

It is also important to note about competitive advantage in platforms that provide services for several clients. A firm is said to have a competitive advantage when implementing a value-creating strategy, not simultaneously being implemented by any current or potential competitors (Barney, J. B., 1991). For this study, the discussion strategy is by building the data analytics capabilities on the firms.

One cold analysis could say the platform itself could play the role of the strategy that creates value. Being a platform that one subscription is enough would not be a competitive advantage since competitors can easily acquire it. As Barney (1986a) wrote, technologies that can be purchased across markets are likely to be imitable and thus not a source of sustained competitive advantage.

However, other researchers demonstrated that, even though the technology can be readily available for different firms, only a few can create a close human-computer interface. Thus, this kind of information processing system may be rare (Christe, 1985; Rasmussen, 1986). It is also a socially complex system and will probably be imperfectly imitable (Barney, J B, 1991). Moreover, it reinforces this study's proposal - where the DCV strategic theory better explains the interaction with the technology, the processes, and the need for human resources and skills. The competitive advantage can be achieved by building capabilities that may not be imitable. The interaction between the technology and the firm does not happen in the same form. Discovering how these process works show its value.

## **2.5 BIG DATA ANALYTICS CAPABILITIES (BDAC)**

The term BDAC (Big Data Analytics Capabilities) was proposed to define the broader range of aspects that big data projects must have to succeed (Mikalef et al., 2018). In the literature, there are different definitions for BDA. However, they have the same essence by defining BDA as a firm's ability to provide insights using data management, infrastructure, and talent to transform a business into a competitive force (Akter et al., 2016a; Kiron et al., 2014).

This research could find few studies about the application of BDAC and its impact on firms. *Big data analytics and firm performance: effects of dynamic capabilities* (Wamba et al., 2017) and *How to improve firm performance using big data analytics capability and business strategy alignment?* (Akter et al., 2016) there are examples where the authors were able to quantify the importance of BDAC for companies. In the first study, the authors also suggest testing the theory in other contexts. Thus, this study seeks to find the relationship between BDAC and the platform's companies understudy and diagnose if it can boost, create, or incentivize BDAC on its clients.

Both studies cited above use the same typologies to model BDAC. For this study, the concepts below will be used as a reference for evaluating the BDAC of the companies. The ideas are also described and used in both studies. An overview of all typology literature can be found in Akter et al., 2016, and it is also presented below.

**Table 2 - BDAC typology overview**

Related Studies	BDA management capability	BDA technology capability	BDA talent capability
Kiron et al. (2014)	Analytics planning, sharing, and coordination, investment, control on analytics as a whole.	Organizational openness, compatibility analytics technology, collaborative use of data (connectivity).	Analytical talent, technical and business knowledge, the organization as a whole effective in disseminating insights.
Davenport et al. (2012)	Analytics management at the core business and operational functions.	Open-source platforms (e.g., Apache Hadoop and cloud-based computing) ensuring connectivity, compatibility, and modularity.	Data scientists or human resource capability
McAfee et al. (2012)	Corporate strategy	IT infrastructure	Skills and knowledge of data scientists
Wixom et al. (2013)	Strategy (e.g., cost, service, price, productivity)	Data (e.g., data model, standard, and control)	People (e.g., capability to use basic reporting and ad-hoc query tools, performance management dashboard applications, customer-facing web portal applications, etc.)
Barton and Court (2012)	Management (ensuring data and models work together).	Data (volume, variety, veracity, etc.) and IT platform.	Talent (e.g., capability to build advanced analytics models for predicting and optimizing outcomes).
Wamba et al. (2015)	Management (planning, investment, and control)	Infrastructure (connectivity compatibility, modularity)	Talent (management, technical, business relational, etc.)
Ransbotham et al. (2015)	Management (planning options, co-ordination between analytical producers and managers, model-based	Infrastructure and processes (machine learning, data management, and information systems) to improve data quality.	Talent (e.g., domain knowledge, statistics, and other technical skills).



	decisions and control)		
--	------------------------	--	--

Source: Akter et al., 2016

The quantitative studies about BDA mostly use the compilation made by Akter et al. (2016). In 2017, Mikalef and his colleagues expanded the typology above, arguing that most studies to date were focused on resources and processes. Still, they did not offer much insight into how firms can develop a strong DBA capability (Gupta & George, 2016). The result is a synthesis built on DCV (Dynamic Capabilities View) foundations in three main pillars that allow firms to develop a BDA capability: tangible resources (e.g., infrastructure, IS, and data), intangible resources (e.g., data-driven culture, governance, social IT/business alignment), and human skills and knowledge (e.g., data analytics knowledge, and managerial skills).

In a more recent study, Mikalef et al. (2019) conducted research based on past studies to create a research agenda over BDAC on firms, and he suggested a framework on how scholars could approach this problem. The proposed framework highlights the importance of factors that pertain to processes, people, technology, organization, and data (Mikalef et al., 2019). The elements are a group of capabilities.

### **3. METHODOLOGY**

Companies that offer big data analytics as a service are still emerging. This environment is even rare when provided for small and mid-sized companies. By looking at existing studies about this subject, it is impossible to find content on specific cases like the mentioned in this study. In this matter, this study can explore the challenges in the initial stages of a business adopting big data analytics in its decision-making process. As also proposed as a research question to be answered, we explored what are the constraining forces, pressures, and enablers while adopting and maturing big data and business analytics (Mikalef et al., 2020).

Having this in mind and exploring the opportunity to have privileged access to a company that offers this service, this research aims to conduct a case study. The case study justifies itself by being a field that lacks existing studies for this specific context. Mechanisms of the big data analytics capabilities can be explored, and the main goal of the methodology is to examine proposals of future testing and research (O'Brien, C., 2007).

By definition, a case study is qualitative research, an approach for exploring and understanding the meaning individuals or groups ascribe to a social or human problem where data is typically collected in the participant's setting, data analysis inductively building from particulars to general themes. The researcher makes interpretations of the meaning of the data (Creswell, 2014).

A case study research explores a program, event, activity, process, or one or more individuals in depth. Cases are bounded by time and activity, and researchers collect detailed information using various data collection procedures over a sustained period (Creswell, 2014). In this case, the purpose is to develop big data analytics capabilities in a company that provides big data analytics service for companies of the real estate market in mid-2019 and the entire 2020 to develop these clients' capabilities.

#### **3.1 THE CASE STUDY**

The case study is called iPPi, a data analytics platform designed for real estate companies. The company behind iPPi is called 4property (formerly 4pm), an Irish company that provides software solutions for real estate companies. The company has around 20 years. So far, their core services were to give full support to the Acquaint CRM, offer website

development and design and data integration between solutions of real estate environments (CRM, websites, public listing web portals, and other third-party solutions).

iPPi, the data analytics platform, is an abbreviation for Independent Property Price Register. The solution was first conceptualized in 2014, with the proposal to be a market analysis and forecast portal for the real estate market. At that time, 4property had all the technical expertise to aggregate and process the data from the different sources available on the market acquired by developing all system integrations and having their property database called Property Drive.

The Irish market has a particularity called Property Price Index (PPR), a government organization responsible for registering all real estate transactions in Ireland. PPR's data was a valuable source to create an analysis for the Irish real estate market. However, the data delivered by this entity lacked information. For the transactions, the address, the value of the trade, and the date were the only data registered. Even the addresses were published lacking information like street names, the house number, number of beds, baths, and size of the property.

To work around the lack of the property price register data's meaning, the 4property owner had the first idea to store all the property price register data and their internal data and information from the publicly listed property listings portal to rebuild complete details of the property transaction. The company could expand the information of property price register matching all the different sources. The result was a complete knowledge about the transaction, property features, and its phase during the period they were for sale, the information from the Ireland census data, geolocation from online geocoding services, and other data aggregation from different sources.

The second phase of iPPi consisted of wrapping up all this data and making it sellable for groups that could use all this information in their business. The first niche of customers that came to mind were real estate agencies, which 4property had already experience and contact.

The first tentative was to create a portal where users could generate market reports for Ireland's different regions and areas. The first try did not call enough interest from the customers. Although they showed interest in the platform, they barely visited the platform after the test.

The second tentative to create a product with data and tool where real estate agents could use a workflow where they could easily find similar properties that they were

prospecting a sale or sold a property and create a valuation for that property based on these found comparables and the property specificities (the company calls it as the appraisal workflow tool). This time this tool experienced more attention so that customers were interested enough to pay for the service.

Nowadays, iPPi is working to improve its user base and user experience. There are several ideas to create value from data. Furthermore, 4property created a product from the data and is willing to add even more value for its customers. The platform was first idealized to be used as a market analysis tool, but it started to be used in other forms and perspectives. This study wants to explore it. Does this service create real value and give competitiveness to its customers? Is this platform in some way contribute to its business customers build big data analytics capabilities?

### **3.2 DATA COLLECTION AND ANALYSIS**

The company focus of the case study offered access to data for this investigation. This access includes documents, access to the platform, employees, and clients.

After the literature review, the first source of information was the analysis of the platform. By having access to the platform, we collected information on the data available, type of reports available, and tools.

The second source of data collection was people. Interviews with the service supplier and clients were done, with a guide to explore the big data analytics capabilities dimensions. It was important to interview technical and management employees from the supplier side to explore the different dimensions. From the client's side, looking to approach a heterogeneity of customers' segments, each of the following types of client employees is interviewed: real estate agencies, investment companies, and government agencies. The interviews were done through Skype, recorded, and all material was transcript to a text document.

The interview guide was based on the study of Mikalef et al. (2019), which evaluates the presence of the major dimensions of the big data analytics capabilities (data, technology, people, organization, and process) by having specific questions for each dimension. There is a set of questions to cover the period before and after the iPPi platform user adoption. They were used to identify the impacts for each BDAC dimension. There were slight modifications in the questions to make it easier to understand our interviewees. Most of them were from small companies without a deep understanding of specific topics, such as technological ones.

The script of the interview is presented in Appendix A and B. All the questions were used as a guide.

An analysis of interviews was then made over the answers' content in a qualitative approach via content analysis. Even though some categories were already available to guide the analysis derived from the literature review's conceptual model, categories were added or removed according to the collected information and the context interpretation.

The other data sources were: documentation, data available are project proposals, market study about analytic (conducted by the supplier in partnership with the University College Dublin), and advertising material (websites, social media posts).

All this information helped prepare for the interviews and fill gaps eventually and reinforce the information provided. Below is a summary of the data collection and analysis protocol. The first column presents the major dimensions of the big data analytics capabilities to be investigated, followed by columns representing the classes in the investigation of the service supplier and service clients. For each cell, we present the data collection sources for each dimension described on the Dimension column.

**Table 3 - Data collection protocol**

<b>Dimension</b>	<b>Data Collection</b>	
	<b>Supplier</b>	<b>Client</b>
<b>Data</b> (Volume, variety, integration, multiple sources)	Documentation, platform analysis, technical interview	Interview
<b>Technology</b> (Infra-structure, tools, know-how)		Interview
<b>People</b> (Management)	Not applicable	Interview with managers
<b>People</b> (Technical)	Interview with managers/technical staff	Interview with technical staff
<b>Organization</b>	Interview with managers	Interview with managers
<b>Process</b>	Interview with technical staff	Interview

Source: The author

In the **data** dimension, the idea was to identify the 7V's (based on the compilation of the 7V's from Mikalef et al., 2017) that defines a big data project. The analysis consisted of using the interview to identify each V's existence, along with the platform's analysis, any documentation available, and advertisement material. The focus of this dimension was the service provider. Even though there was less focus on checking it in the platform's clients, there were also questions investigating the maturity of data that each client held in the interview guide. If there is any sign of a more progressive approach to maintaining data, further questions could be made to explore this issue further.

The purpose of the **technology** dimension was to discover the technologies that are around the data held by the company. For this dimension, the focus was on interviewing the skilled technical staff of the companies. It could also be used as auxiliary material from any documentation and platform analysis. From the interview, it was necessary to identify if the company used cloud solutions, technologies for processing large amounts of data, the tools to operate and visualized the data, the existence or use of databases specific for dealing with big data, and the in-house expertise of the company on managing these technologies (based on the BDA technology capability from the authors' Davenport et al., 2012, Kiron et al., 2014, McAfee et al., 2012 and Wamba et al., 2015). The technology was another dimension with a focus on the service supplier. The questions made for the clients only investigated technologies for the most basic use of data. However, if it was noted a higher level of maturity in the company, the same investigations about the technologies were viable as having an internal IT team.

The **people** dimension investigated two aspects of the company. First, the goal was to identify managers' vision about big data, seeking signs of incentive from the managers, and the awareness about the business value (top management support, according to Mikalef et al., 2017) using data analytics. At the technical level, the second aspect was checking if the companies had skilled staff to deal with data and or the technology stack and programs of incentives for training about data analytics topics from the company. For this dimension, the technical aspects were concentrated on the service supplier. For both clients and the service supplier, the investigation about the management aspect occurred.

The **organization** dimension intended to identify the signs of formal strategy, business planning, investment in big data projects, and awareness existence about costs, productivity, price, business-IT alignment, and data-driven culture (based on the BDA management

capability and the authors' Davenport et al., 2012; Kiron et al., 2014; Wamba et al., 2015; Wixom et al., 2013). This dimension was analyzed with both clients and the service provider.

The last dimension was the **process**, which was to identify signs of existing processes involving data. For example, any formal processes and management control over the data (access, add, delete), uses of data on routines at management decision-making or operational level (based on the operational capabilities of BDAC from Mikalef et al., 2017). Both clients and the service supplier were the subjects of this investigation. From the service supplier lens, the focus was on processes around the data management and the service provider's clients. The analysis was on the use of data at routines of the company.

The analysis for each dimension was made by analyzing all the material and checking its aspects in the service provider's clients. This research did not use any tools or a framework for the analysis of the content itself.

**3.3 OVERVIEW OF THE SUBJECTS**

Eight companies participated in this study - the case studied and the other seven companies from the data analytics service providers. For each client, there was one interview, usually managers or a real estate agent. For the service provider, there were two managers interviewed, the CTO and the CEO. The interviews were conducted during December 2020 and January 2021. All the other material used for the analysis was collected through the year 2020.

This report did not provide client's and interviewees' real names to guarantee the participants' anonymity. Each company and participant were described using a generic descriptor and a code letter. A brief overview of each company is presented with a description of the interviewee's role in the research interviews in table 4.

The company's numbers are related to the number of transactions of residential sales and residential rent the company owns, along with the descriptions. The purpose of these numbers is to use them as a form of comparison of their size. These numbers are only available for the clients.

**Table 4 - Companies represented in the Interviews**

<b>Company A – The data analytics service provider</b>	Interviewee 1: CEO, responsible for all strategic decisions and products of the company.
--	--

	<p>Interviewee 2: CTO. Responsible for all technical, strategic decisions of the company.</p> <p>Description: Company A is the company of being the case of study. There is already a complete description of the company in section 3.2.</p>
<p><b>Company B – Real estate agency with multiple branches in Ireland</b></p>	<p>Interviewee: Director of one branch of the company.</p> <p>Description: Company B is a traditional Irish real estate company. The company operates through multiple branches around the country. It has around 80 branded offices in the major cities and towns of the country.</p> <p>The company offers all services around the property market: Sales, auctions, management and letting services, valuations, asset strategy and optimization, planning and development advice, etc.</p> <p>Company B has its own IT department that provides IT services for all branches.</p> <p>The performance of company B in the year 2020 was 3,713 properties sold, which the median price of the property sold was 316,000 Euros, and 1,364 rented properties with a median rental price of 1,200 Euros (4property, 2020).</p>
<p><b>Company C – International real estate agency - franchise</b></p>	<p>Interviewee: Franchised owner (CEO) of the branch in Ireland.</p> <p>Description: Company C is a multinational company. The brand is the oldest real estate franchise system that originated in North America. They call themselves the first original real estate “startup”.</p> <p>The company franchises enjoy complete support from its franchisee. They have available to use a complete platform that offers property management systems, data management, mobile applications, CRM, and more.</p> <p>The company offers a range of services around real estate transactions.</p> <p>Company C's performance in the year 2020 was 42 properties sold, which the median price of the property sold was 260,000 Euros, and 97 rented properties with a median rental price of 1,450 Euros (4property, 2020).</p>
<p><b>Company D – Local real estate agency</b></p>	<p>Interviewee: Director of the company.</p> <p>Description: The company D is a local real estate agency long-established, 3<sup>rd</sup> generation family firm that operates in only one county of Ireland. Company D operates in all aspects of sales, rentals of residential, commercial, and agricultural properties.</p>



	<p>Company D's performance in the year 2020 was 41 properties sold, which the median price of the property sold was 165,000 Euros, and 17 rented properties with a median rental price of 950 Euros (4property, 2020).</p>
<p><b>Company E – International real estate agency - franchise</b></p>	<p>Interviewee: Regional operating principal</p> <p>Description: Company E is a franchise of a global real estate company that operates in a model that uses more technology, and as they say, they are moving away “from the high street branches”, offering services through technology and super personal service. They offer a platform for agents that are part of the agency, giving them more flexibility and freedom to work. The agents' relationship is not through an employment contract; instead, the agents' payment is a percentage of the property transactions. Though the company E, agents have access to CRM, personal websites, and mobile applications.</p> <p>The brand has its own IT team that develops most of the company’s software solutions.</p> <p>Company E's performance in the year 2020 was 663 properties sold, which the median price of the property sold was 400,000 Euros, and 47 rented properties with a median rental price of 1,700 Euros (4property, 2020).</p>
<p><b>Company F – Local real estate agency</b></p>	<p>Interviewee: Sales development specialist</p> <p>Description: The company F is a small real estate agency located in Dublin. The company operates mainly in specific regions of Dublin, the capital of Ireland. The company's main operation is in residential sales and the residential letting property market. It is a small company operated with three employees.</p> <p>It is a small company that is still using non-standard platforms for operating its business.</p> <p>Company F's performance in the year 2020 was 49 properties sold, which the median price of the property sold was 300,000 Euros, and four rented properties with a median rental price of 2,400 Euros (4property, 2020).</p>

<p><b>Company G – Local real estate agency</b></p>	<p>Interviewee: Manager director</p> <p>Description: The company G is another small real estate agency located in Dublin. The company covers few regions from Dublin, and its main operations are residential sales and residential letting property market.</p> <p>Even though it is a small company, company G runs its business using mature cloud platforms to support its activities, as CRM is focused on the property market.</p> <p>Company G's performance in the year 2020 was four properties sold, which the median price of the property sold was 610,000 Euros, and 63 rented properties with a median rental price of 2,000 Euros (4property, 2020).</p>
<p><b>Company H – Local real estate consultancy company</b></p>	<p>Interviewee: Company founder/owner</p> <p>Description: The company H is a local real estate consultancy company with a strong focus on commercial properties. The company activities include advice on acquisitions, valuations for large schemes, and commercial property valuations.</p> <p>The company has three employees and is in Dublin. They use technology tools like CRM, but as noted by the interviewee, because of their work, as they have few clients, the current cloud tools are not relevant for their business since they focus on more sales activities.</p> <p>Company H's performance in the year 2020 was one property sold. The property's price was 340,000 Euros, and seven residential and two commercial rented properties with a median rental price of 2,000 Euros (4property, 2020). It is important to note that the main area of work of company H is consultancy. The numbers presented here can not be used as a form of performance comparison with the other clients.</p>

Source: The author

### 3.4 GENERAL OVERVIEW OF THE COMPANIES

To have a better overview of the companies when comparing them with characteristics relevant for this study, a summary below highlights their differences.

**Table 5 - General overview of the company’s participants of this study**

Attribute	Relevant profiles
<b>Business model</b>	The business model varies from the agency with a single office from an agency with multiple offices to offices from a franchise

	model.
<b>Market segment</b>	The companies were not equal in the market they acted. Some were strong in the residential sale, other ones in the residential renting market. The transactions' price segment also has a significant variance, where some companies were mostly working with low price rental and sales properties, and others acting with a higher ranged price. The companies' location was also mixed, where there were companies located in countryside regions, and some located in central regions.
<b>Technology adoption</b>	Most of the companies were already adopting mature technologies in their process. But on different levels. One company demonstrated a mature use, even having developed its tools (The franchisee). 2 of them having their own IT teams. Only one company demonstrated a low use of technology in their process (The company focused on consulting). The most used tool is CRM.
<b>Number of transactions</b>	<p><b>Sales:</b></p> <p>Smallest: 1 sale</p> <p>Largest: 3,713 sales</p> <p>Median: 42 sales</p> <p><b>Rental:</b></p> <p>Smallest: 4 rent</p> <p>Largest: 1,364 rents</p> <p>Median: 47 rents</p>

Source: The author

By checking the information above is noticed that the profile of the company's part of this study has a significant variation. The non-linearity of subjects is important since the views and generates insights from different perspectives.

## 4. ANALYSIS AND DISCUSSION

Data analysis is presented in sections starting with an overview of the product, followed by an analysis of each dimension based on the big data analytics capabilities framework summarized in table 2. All analysis is a compilation of interviews and the collected document material.

### 4.1 OVERVIEW OF THE SERVICE PROVIDER

The service provider platform identifies itself as a big data analytics platform. The core data is the Property Price Register, a government database that publishes all property transactions in Ireland. Other sources augment this data. Their sources are:

- **Public property listings portal** - portals that publish properties for sales. Well known as “Property classifieds”, they are extremely popular channels to search and publish properties for sale. iPPi works by crawling these classifieds every day, maintaining the property's information and its features updated.
- **Census Data** - Census is the procedure of systematically acquiring and recording information about a given population's members. Like most countries, the Ireland government also runs its census and makes its result public. This information is important for iPPi because it can use the definitions of areas and cross-data to generate statistics at a small regional level. The census is run every five years. So, this information does not need to be updated so often.
- **Google and Microsoft APIs** - Google and Microsoft offer several services for consulting data online. They are valuable for the service being provided, where some APIs offer detailed information about geolocation and addresses. Google maps, for example, has several services to decode raw addresses. By doing so, they can return from raw address information of its geocode location (latitude and longitude) and the detailed information about nearby amenities.
- **4Property data** - The company itself has valuable data about its customers and their public listings of properties. Although they are limited compared to the public online classifieds, they are a source to reinforce its dataset.

These are the main data sources that feed the platform. There is also another secondary dataset and plan to insert new data sources into the iPPi data environment. A sample of aggregated data of a property is presented in Figure 2.

Figure 2 - Sample of aggregated data of a property

**16 Fishermans Wharf, Washington St, Cork** ...

2 🏠 | 1 🛏️ | 60 m² | Apartment BER B2 There are 0 open challenges! Apartment

Sherry FitzGerald are delighted to offer 16 Fishermans Wharf a bright and spacious second floor river side apartment located in the heart of Cork city centre. The property comes to the market in very good condition and enjoys a very sunny south and west facing aspect overlooking the south channel of The River Lee. This quiet apartment would hold appeal for someone seeking an ideal investment or a first-time home with the convenience of modern city living.

Accommodation includes an open plan kitchen/dining/living room with sliding door onto a private balcony onto the River Lee, two double bedrooms and bathroom.

Positioned within a few minutes walk of UCC, Bons Secours Hospital, Mercy Hospital and St Patrick Street location is excellent.

A viewing is a must to fully appreciate the qualities this apartment has to offer. Don't delay arrange a viewing today.  
Entrance Hall 4.3m x 1.0m. Private hallway with plaster coving, ceiling rose and coat hanging area.

Kitchen / Living / Dining Room 4.5m x 4.3m + 4.7m x 3.3m. Generously proportioned living space enjoying a sunny south and west aspect with sliding door to balcony overlooking the River Lee. Fully fitted kitchen with eye and floor level units with all appliances. Wall mounted shelving unit. Tiled over counter splashback.

Bedroom 1 3.4m x 2.9m. Spacious double bedroom with window overlooking the River Lee. Built-in wardrobe.

Bedroom 2 3.5m x 3.0m. Generous double bedroom with built-in wardrobe.

Bathroom 3.4m x 1.6m. Three-piece bathroom suite with electric shower fitting. Tiled floor and part tiled walls. Hot press store.

**DIRECTIONS:**  
From The Grand Parade turn left onto Washington Street. Go through the traffic lights at The Courthouse. Take the first left turning at Costigan's Pub leading onto Hanover Street and turn right. Fishermans Wharf is along on the left.

Map showing location near West Link Business Park, City North Business Park, Glensprings, Rockgrove, and Blackpool Shopping Centre.

Price: **€ 250,000**  
€ 4,167 / m²  
Advertised Price: € 250,000 **10.0%**  
Current estimated market price inclusive of VAT:  
**10.0% € 250,000 ± € 9,547 (95% of confidence)**

**311** DAYS TO SELL  
Launch Date: **14/02/20** First seen  
Sale Agreed: **30/11/20**  
Sale Date: **22/12/20**

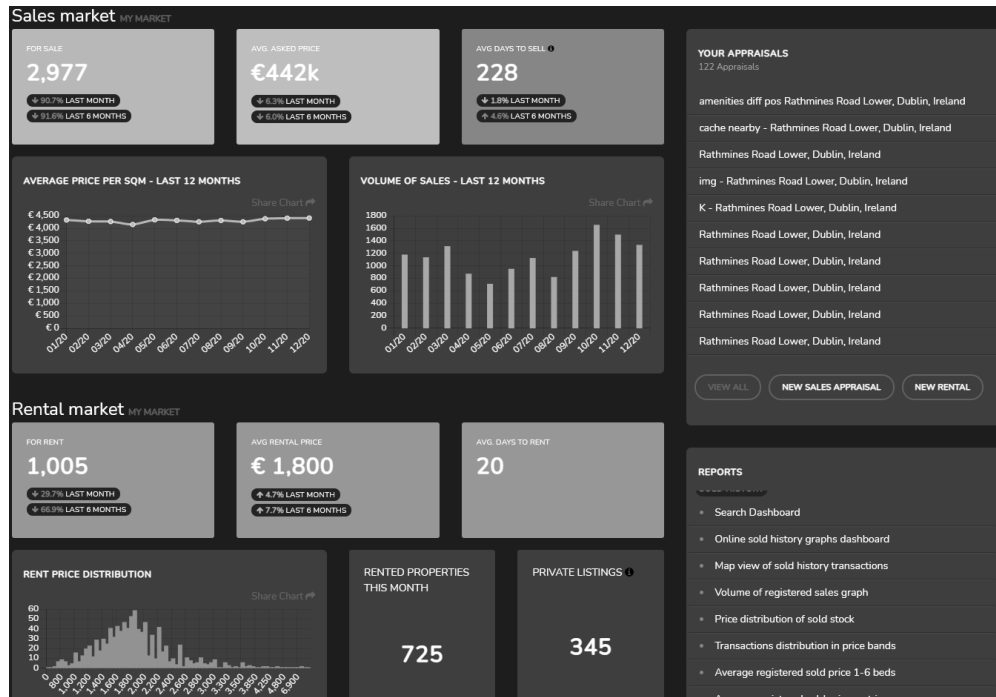
Gallery of photos showing the apartment and its view of the River Lee.

Source: iPPi platform

Using the 7 V's of the big data definition, we can validate the affirmation that iPPi is a big data analytics platform for the real estate market.

**Volume** – As we can observe in Figure 4, the data being collected is considerably big. The platform does not concentrate only on the property addresses and the property details information. The platform can aggregate other sources, like census data, to generate different types of information.

**Figure 3 - Platform dashboard with a compilation of the data available**



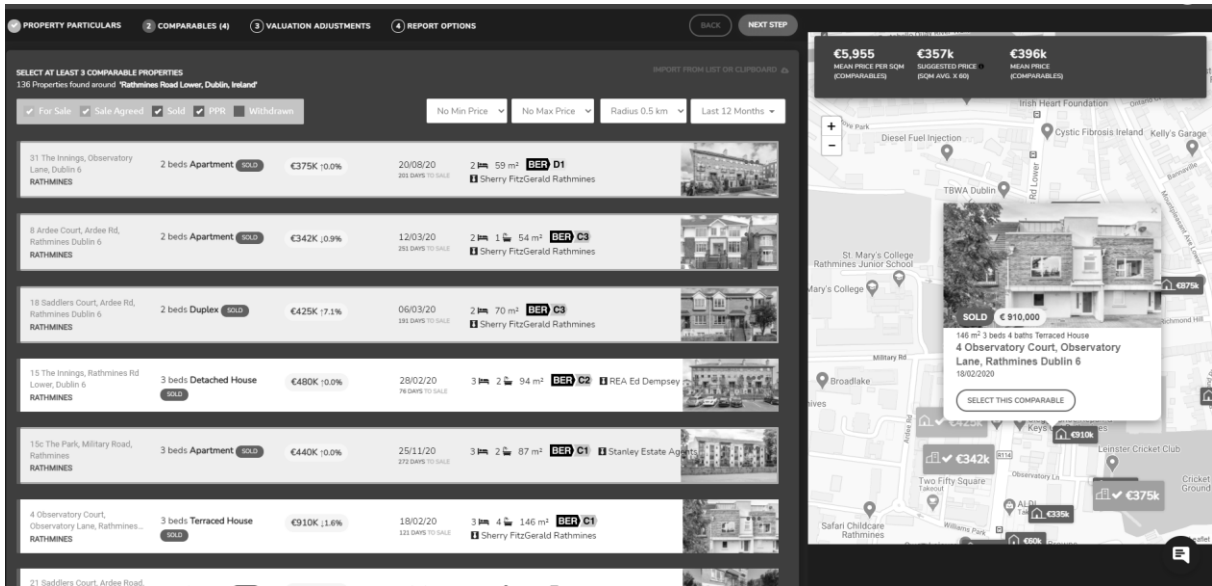
Source: iPPi platform

**Velocity** - Although most of the dataset does not need to be collected and analyzed in real-time, it is updated and processed daily. The velocity for the current purposes meets its clients' needs. Nevertheless, there is room for improvement. We can cite the possibility to use real data about the economy, market trends, and social media.

**Veracity** - The data comes from different sources. The Property Price Register and the Census data can be considered reliable since the government publishes them. The other validity of the other sources is also crossed to reassure its validity.

**Value** - The value generated by iPPi can be considered worthy from the moment clients are seeking it to do property valuations, as Figure 5 shows. The valuation of property follows market trends and consumer behavior. The platform is providing all this information. It can be considered the principal value, but other types of use are expected. There is the expectation that real estate agencies can also use the charts and the information provided by iPPi to make strategic decisions at the company's management level.

**Figure 4 - Tool for property valuations**



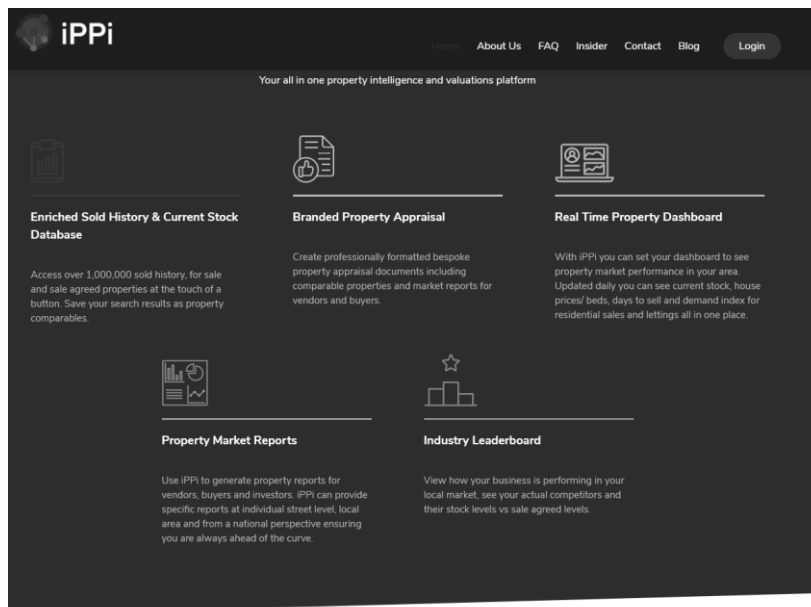
Source: iPPi platform

**Variability** - iPPi offers several reports about the real estate market. It does not mean the user is restricted to this information. Other several analyses can be done over the iPPi data by third-party tools and developed in-house. The level of variability is high since all data can be crossed using geolocation.

**Visualization** - iPPi tries to fit well this V is offering a set of reports over its data.

In sum, iPPi shows to cover all V's from the 7 V's definition of big data. We can conclude that the iPPi platform is a genuine big data project, as they announce (see Figure 5). Although it covers all the 7 V's, they are covered in different levels and still can improve in some aspects.

**Figure 5 - Advertisement of the data analytics platform**



Source: iPPi platform

## 4.2 ANALYSIS OF THE DATA CAPABILITIES

The first dimension of the protocol, based on the Mikalef et al. (2019) model, is the data capabilities. The service provider has already access to large (the volume characteristic George et al., 2016), unstructured, and fast-moving (the velocity characteristic, George et al., 2016) data and the degree to which it integrates its internal data and external (the variety characteristic, George et al., 2016).

There are several processes to keep the code's quality (the veracity characteristic, Wamba et al., 2015) in the service provider. It includes the ability of the clients to challenge the data. Workflows to ensure the locations and avoid conflicting information. Data as evidenced by the manager interview and the previous overview of the service.

The data was under constant changes and updates. One example is the matching data and the constant updates in several indexes of the property data information. One example of these updates is the need to compute a specific property's estimated current price (the variability characteristic, Seddon and Currie, 2017). The service provider offers several ways to visualize the data. The main one is the interface that clients can access through their subscription. There are also several visualization tools for internal use (the visualization characteristic, Seddon and Currie, 2017).

“We are also acquiring data from multiple sources, at through the industry, and increasing the government's register of transactions. What we are doing then is



correlating all these multiple data sets. And we have a million records effectively. And we were analyzing and correlating and these transactions that we're looking at in terms of doing further and correlations and possibly perform more data science in terms of ml and AI.”

- *Interviewee CEO Service Provider*

From the client's perspective, it was evidenced different levels from ownership and access to data, and how they view data as a resource. Most of the companies reported the same as company H. The interviewee described his company doing only basic data usage, using primitive technologies to store data and manage them before using the platform.

“Not a huge amount of data is held by our organization. And, you know, we are kind of about a boutique kind of firm. So, we only handle a handful of clients. And in terms of data, we don't deliberately collect data.”

- *Interviewee Company H*

Most companies demonstrated that they keep most of the data held by them through tools like CRM. And it was not a deliberated strategy to store the data as an important resource.

“Obviously, we have all our tenancy details in terms of, you know, tenants' names, phone numbers, email addresses, and PPS numbers from people who are actually who rent from us and we look after on a daily basis. [...] it is basically a CRM system for it for the office. covers everything from our lettings, to our management, to fears to action to cover everything up.”

- *Interviewee Company B*

“We have a CRM. But we also use data from property week<sup>1</sup>, from iPPi be from the PRA<sup>2</sup>. And from and from the commercial leases register. Also, by old fashioned calling other agents.”

- *Interviewee Company C*

“Everything we hold (Data), we have a server here and everything is in house, we have no online backups, we have nothing, nothing, nothing online, everything is held here. Obviously, we went through a CRM, we have details, contact numbers, addresses, emails, for all of the vendors and the potential purchasers of properties. And then data, like all that data stored through the CRM, any correspondence to and from any of those clients is stored in the CRM, and then every document that we produce here is stored on our server.”

- *Interviewee Company D*

---

<sup>1</sup> Property week is a Irish company that provides real estate data transactions that are manually collected.

<sup>2</sup> The Property Registration Authority (PRA) is the State organization responsible for the registration of property transactions in Ireland.

“So basically, the data we hold will be obviously you know, names, numbers, and email addresses and then property I suppose, like whether someone is selling a property would also hold that nation will be also hold buyer’s information for buyer matching.”

- Interviewee Company F

In terms of data, the service provider's clients were limited by their data before using the platform. Alone they could not have a big volume, variety, and the possibility to integrate data with multiple sources. Using the data analytics platform, we can conclude that they could access a huge volume of data without investing in technology, people, and data management. The service provider fulfilled this capability.

“Prior to using iPPi, we used our own database to see properties that we had sold. So, if I am doing a valuation of your house, I might have sold houses in that area before. So, I'd use my own database, look back and see what other houses I have sold in that area before, **but that's limited because of limited only two houses that we have sold.** Whereas with iPPi, I'm seeing every house that came for sale in that area.”

- Interviewee Company D

“We use the iPPi and we find that extremely useful like we've only started using it really probably this year. And since we started to use it as a database to find comparables for valuations.”

- Interviewee Company F

In terms of data capability, we can identify all its aspects in the service provider. In terms of owning big data capabilities, all seven V’s were found. The last V that was not described yet was the value. However, the value can be seen in the platform itself and the different ways the clients were extracting value from it and will be described below.

### 4.3 ANALYSIS OF THE TECHNOLOGY CAPABILITIES

The analysis of the technology capability is based on the service provider. The reason is that most of the clients reported that they do not have internal IT teams for developing and managing their infrastructure – they are outsourcing this resource. They rely on cloud solutions or only having only one person with a technical role. Two companies reported having their own IT team, but they were inserted in the franchisee's business model. The IT team was for the franchise, and it was not under the franchisee.

“He is our tech guy. He does. He looks after all of that (the IT).”

- Interviewee Company E

The service provider reported a level of maturity in terms of technology-focused for big data analytics. The interviewees pointed out all the topics, such as open-source platforms,

cloud solutions focusing on data processing, data storage with NoSQL, and relational databases.

“We are currently exploring open-source software for it. So, there are a number of platforms we are examining. Now there is a platform called Cloudera SDK<sup>3</sup>. So that's kind of open. It is based on open-source projects. It is a conglomerate of open-source projects. We are also looking at a technology called confluent.io<sup>4</sup> in the streaming domain. [...] And ways we have used inside of Elasticsearch<sup>5</sup>, we use the Kibana<sup>6</sup>, regarding being able to just query data, and we are using them and some bespoke kind of graph, graphical kind of frameworks to actually be able to analyze and display some of the data on the front end.”

- *Interviewee CEO Service Provider*

It is also interesting to note that the company has awareness about monitoring the new technologies, testing them, and reproducing the market standards for big data processing. Moreover, the entry cost is getting smaller, according to the interviewee.

“I feel that any company that can handle the technology appropriately, and the cost of entry is becoming less, and they are becoming more and more standardized. So that's when companies, like 4property can exploit it because that means it's not high technology anymore.”

- *Interviewee CTO Service Provider*

We can also conclude that the service provider can replace the development of the companies' needed technologies, giving them fast access to explore the outcome of big data. Even though clients do not have the technological infrastructure or sufficient technical support to handle big data analytics, they can access this as a service. Their partner (the platform) provides them databases, cloud, and other technological resources for big data analytics. Moreover, the technology aspect is easy to identify once it is a tangible resource (Mikalef et al., 2019).

#### 4.4 ANALYSIS OF THE PEOPLE CAPABILITIES

In the people aspect, there are two different categories to be analyzed. This one is related to technical human resources. For this aspect, the biggest focus is the analysis of the service provider interview. There was no reason to get into this on the service provider's clients since it was clear that they did not have IT teams prepared to deal with the data

---

<sup>3</sup> Cloudera SDK is an enterprise data development platform.

<sup>4</sup> Confluent.io is a company with a main product for data processing.

<sup>5</sup> Elasticsearch is a no-sql open-source database optimized for search.

<sup>6</sup> Kibana is a tool for data visualization.

subject. The second category is the general staff skills to make use of big data in decision making. Data analytics skills are fundamental for successfully executing BDA and further developing a data-driven culture in the clients. Aspects of organizational culture are going to be approached in the analysis of the organization's capabilities.

Starting with the service provider, the company demonstrated human resources aspects in terms of the big data analytics capabilities from the interviews. They are skilled technical staff, training, technical and managerial knowledge.

“We provide the data on its training to our own employees. [...] the knowledge that the people don't have, and they need to seek that out and do courses online.”

- *Interviewee CEO Service Provider*

“You know, it's all good and well and good for from a technical perspective to look at the tools. [...] but we need a formal R&D practice.”

- *Interviewee CTO Service Provider*

However, both interviewees demonstrated some concern about skilled people with data science skills. According to them, they are progressing with the infra-structure skills that are the base of the platform. Nevertheless, people with data science skills are difficult to find in the market. Furthermore, they are trying to attack this issue.

“I think that there's a shortage probably of people available with the really good skills in data.

- *Interviewee CEO Service Provider*

“You need to have a proper data analytics team with trained data science experts. We do not have that. So how are we trying to attack this [...] Now one of the initiatives that has been happening [...], I have got connections with academic community in Ireland, because we are placed in the Innovation Center. We are trying to exploit some of University ecosystem to enable or simulate some of the benefits that we can get as a startup. So that is one of the so even there the problem we are facing is we still have not got a good data science practice.

- *Interviewee CTO Service Provider*

The focus of analysis is on analytical skills for using data in the decision-making process for the clients. Do the managerial staff or the operational staff has skills or training in the topic? When questioning if the companies provided training on data analytics topics, most of the answers were mostly negative.

“No, I don't think so.”

- *Interviewee Company B*

“We provide training certainly on the CRM that we have on data analytics, not really.”

- *Interviewee Company D*

Some interviewees reported having only training sessions about the topic because it was provided by the data analytics platform's service provider even though it was more usage of the platform training.

“I got some training videos for the equations, and I find them very useful.”

- *Interviewee Company F*

“Yes (the company provided training) [...] I suppose the iPPi made more pronounced.”

- *Interviewee Company G*

However, most interviewees said that data analytic skills were important for their market when questioned.

“It gives your precision.”

- *Interviewee Company B*

“For valuation, it is very important, very important, you know, because you probe for value. [...]”

- *Interviewee Company C*

Analytics capabilities can vary from simple uses of dashboards to extract data and perform some analysis to data science capabilities related to the more sophisticated data exploration. The clients interviewed seem to be practicing the simple type of analysis.

In sum, regarding people capabilities, it was found that the service provider was able to fulfill the need for technical skills. However, there are still missing pure data science skilled people responsible to constantly analyze and generate insights from all the data in different ways. This persona could be on both sides. It could be present only on the client's side, where this person was responsible for extracting value from the raw data. The service provider was only responsible for maintaining and feeding the data, with this arrangement could bring the ability to build differentiation for the company since all insights or value extracted from it would be aimed only at that client. Another arrangement is having a data scientist on the service provider, extracting and creating insights that could benefit all the clients. In this arrangement, there is a loss of the potential to build differentiation on the clients, but it could back competitiveness against bigger or more mature companies in the BDA field. There is still room for improving people's capabilities regarding data science capabilities on both sides.

#### 4.5 ANALYSIS OF THE PROCESS CAPABILITIES

This analysis intends to discover the formal process or procedures the companies may have when dealing with the data to generate insight from it. Usually, these processes are related to data governance, including data migration, data retention, cost allocation, analytic data procedures, and access rights (Mikalef et al., 2019). This aspect was only enquired to the service provider's clients, as they are the part that should benefit most of the data service.

Starting with data policies, like backup policies, access policies, most companies show that they naturally rely on cloud services. These policies were delegates by the providers of these services.

“Yes, they take care of security. (the CRM cloud provider).”

- Interviewee Company B

In terms of using the data, the interviewees' main description was using the data for doing valuation of properties. The platform became an important source of data for this valuation. It saves a reasonable amount of time in accessing and generating these reports, which changed the procedure.

“So, we use it, you know, for like, mainly us really I do a lot of valuation work a lot of a lot of bank valuations for mortgages [...] it saves me an awful lot of time.”

- Interviewee Company B

“We are more interested in the valuation. (when the interviewee asked about how the use of the data in the company is)”

- Interviewee Company C

“we're using it for valuation at the moment, but I know that there's a there's some good market insights in terms of basically, you know, pricing or properties per square foot, I know that there's a lot more to comity, so we're excited for that. But at the moment, it is just mainly be used for comparables [...] So we're looking forward to the future with it to see how it could help our business.”

- Interviewee Company F

Only one interviewee demonstrated to use the platform beyond the operational valuations, which was the interviewee of company E, demonstrating a use for competitive intelligence analysis.

“I use iPPi just for keeping an eye on where we are compared to our competitors. [...] just to keep an eye on where we are at, where we are how we are, you know, where we are doing well, and where we're not doing so.”

- Interviewee Company E

There is an evident adoption of a formal process of doing property valuations and indirect data management processes (backup policies, retention policies, security) after the

platform's subscription. However, as mentioned before, these were standardized processes created and inherited from the platform's service provider. Nevertheless, it was not identified the development of any formal process beyond the service provider's ones.

It is important to bring the importance of this capability. It represents how capable the firm is to effectively coordinate activities, resources, and tasks around this large volume of information (Mikalef et al., 2019, Ravichandran and Lertwongsatien, 2015). Beyond a well-established workflow, the idea is that the service provider could influence its clients to build this coordination. The capabilities below can provide insights into how it can occur.

#### 4.6 ANALYSIS OF THE ORGANIZATION CAPABILITIES

Starting with the service provider's analysis, the evidence that the managers had a formal strategy, planning, and investments for big data projects is clear. It can be identified in the passages below.

“[...] And then we also build data systems and data analytics on analyzing real estate data, which is coming from that network. And then we are looking to that, in terms of the future of what we can do with that data.”

[...]

“And we are also then acquiring data from multiple sources, at through the industry. And increasing the government's register of transactions. [...] we're looking at in terms of doing further and correlations and possibly perform more data science in terms of ml and AI.”

[...]

“The acquisition of data is a is a major part of our strategy. And the more data or the more correlated data we have, the more we can augment against, and you know, what is there and will help then in terms of our customers and ourselves to make better decisions.”

- Interviewee CEO Service Provider

“[...] primarily as a valuation platform as I said before, and secondly, as a set of tools for realtors to promote properties and do their business promotions inside the market.

[...]

“My focus area now is to look at that, you know, to understand how much can be how much more data can be consumed? Or how much more effectively consumed? Can we consume the data?”

[...]

“the analytics that is expected from the market today in the kind of ecosystems that we have going to its near real time, you know, so the ability to provide fresh insight is really a plus find, especially [...] I mean, given the current situation of COVID, there is

an increasing migration of many real-world market spaces moving to the virtual equivalence, and that process is accelerated in the last one year

[...]

“But we are not really exploiting the customer information today. That's one of the key goals”.

[...]

“So that is where we want to head and the only way to do that is, obviously use distributed computing, because that allows us to do this economically at scale.”

- *Interviewee CTO Service Provider*

In all conversations are perceived aspects of the strategy's alignment, planning, awareness about costs, price, and the project's returned value through the companies' subscription and the future ambitions to expand the platform. The value is also recognized when the interviewee explains the potential returns their clients can achieve with the platform.

The interviews with clients have not identified any sign of a formal strategy for executing big data projects until the platform's subscription. In general, the interviewees demonstrated some interest level in the subject and provided generic positive comments about the theme.

“I think it's getting more and more important.”

- *Interviewee Company F*

“Very important is the straight answer to be able to understand data [...] where the market is going to [...] what sort of trends are [...] For example, let's say the market is moving for like big houses, or houses that are far away from the city centers.”

- *Interviewee Company G*

“it's going to be everything. [...] So that but data analytics and all that, working on that figuring out what the who are the right people to contact, that's going to be the future.”

- *Interviewee Company E*

The organization's capabilities are a set of intangible characteristics (Mikalef et al., 2017) that are more complex to identify. It is closely related to the company's capacity to build a data-driven culture and how the organization can adapt to changes in its environment. Even though most of the companies did not have a clear data-driven culture, the step towards subscribing to a data analytics platform can demonstrate a movement in its direction. It is expected that companies with data-driven culture consider data as a tangible asset (Mikalef et al., 2019).



## 5. CONCLUSION

How big data analytics platforms influence and build capabilities for big data analytics at SMEs? Although the study's case is a service that is still beginning, this study can conclude that this type of service can bring positive benefits for its clients. By starting with the access to the data, it is noticeable how impactful it is to have a platform that offers high-quality data for its users. Most clients did not have access or structure to build with the same quality data warehouse to store, process, and maintain data from multiple sources. It is not affordable for SMEs. In the case of the service provider, which offered other range of products that potentially increase the ownership of data from multiple clients, also offered a CRM for its clients, access to non-sensitive information from multiple users can build exclusive and rich data. There is an opportunity to build competitiveness for SME's clients of the service provider, building resources that usually are exclusive for big companies. If these resources are applied strategically, probably these companies can develop some competitive advantages.

In technology capability, the service provider has been up to date with the most recent technologies and techniques to process data. In the long term, maintaining the relationship between the employees and the technology can improve the accessibility of tools and advanced infrastructure to compute data.

Both parts, the service provider and the clients, demonstrated to lack of skilled professionals on data analysis (data scientists). Even though the platform's purpose is to provide data, having data scientist professionals benefit the platform in terms of the data's orientation enrichens and possibly creates standardized reports and tools for specific cases. The access of data is not enough if it is not explored properly. The clients also reported a focus of use on operational tasks as an intense use on creating valuations. There was no formal approach to training employees about big data and data analysis. However, some users evidenced that the platform was the only access they had for formal data analytics training. The training can also be one aspect of data analytics platforms. Moreover, the differentiation can be built upon the internal knowledge and how it made the interaction between people and the technology, as initially defended by Christie, 1985 and Rasmussen, 1986. The adoption of the platform signaled a start of an organizational movement towards DBAC.

This study detected the evidence of positive influences of big data analytics capabilities in SME's by providing access to data, breaking technical barriers and costs, facilitating the understanding, and, in some level, training around data analysis and big data

through a big data analytics platform being offered as a service. There was value-creating on speeding up the process of evaluating properties (which is an operational work). Nevertheless, there is still a vast amount of value to be explored in other fields that depend on companies building internal skills to explore them and create a data-driven culture.

Sub subscribing to the data analytics service is evidence of seeing data as something worth the investment. Moreover, a data-driven culture determines the extent to which organizational decisions are made based on the extracted insight (Gupta & George, 2016). However, it is not clear how the companies will further develop data in most of their decisions. At least for the clients' use cases, they were more oriented to operational uses. There was a more standardized workflow tool (the property valuation), the adoption demonstrated to be solid. An important aspect of companies with strong data-driven culture is relational practices concerned with the formalized links between employees of the technical and business sides. This link encapsulates practices and means of knowledge sharing, education, training, and strategic planning (Koopers et al., 2011; Mikalef et al., 2019). Data analytics can play a role in that. However, until now, it did not strategically create a real data-driven culture. The service provider can play the role of bringing important capabilities. Building differentiation and changing the culture is necessary to advance the client's maturity use of data, and probably it will come with time. The establishment of a data-driven culture is related to BDAC that needs to be consistently developed in all its perspective to achieve a high level of maturity.

This project's practical contributions can be directed to companies willing to start a data analytics service. The technological aspects surrounding this kind of service are not the most important part of this context. It is important to explore how your service can influence the build of a data-driven culture in its clients so they can extract the real value of big data. The development of specific tools, workflows, and reports can help with the clients' involvement with the service, but only a small part of this equation. From the client's perspective, it is important to cite that the technology itself will not replace the need to build the mindset of making decisions based on the data. There is a huge potential in a data analytics service. However, its value will only be seen with the investments from both sides.

This study's theoretical contribution is to provide one case study of the topic in a specific context. It is important to observe that different contexts will have different challenges. This study's analysis can seed future studies and theories around this subject that

were not covered before. Most studies only approached big companies with the vast availability of resources. In contrast, this study tries to explore how to bring BDAC to SME's.

## **5.1 LIMITATIONS**

This kind of research offers possibilities to understand different nuances of a specific context. However, it is important to expose its limitations. First, the study was conducted during the event of the global covid-19 pandemic. This event can modify the companies' usual routines and the profile of the customers of the service.

Service maturity can also have different impacts on its customers. Studies conducted in companies with a higher maturity level may provide different nuances of its users' impact.

Another important point to highlight is that clients are originated from a very conservative type of market - the real estate market. Most of the users were in the market for more than 20 years, and it is expected they to be less willing to adopt new technologies. Different markets are likely to show different openness and appetite for adopting new technologies.

## **5.2 FUTURE STUDIES**

The big data analytics topics are recent, and contexts to explore how these capabilities are acquired and how they impact business can be sources of a wide variety of studies. As a recent phenomenon, it is important to be aware those different arrangements be selective on contexts that are more representative and impactful in terms of research and contribution. Having this in mind, the author of this study suggests the following subjects for future studies in the big data analytics field:

- Explore cases where big data analytics services are being provided for different markets, particularly markets distant from the real estate market's characteristics.
- Study cases where clients of big data analytics services already adopt data analytics routines for their decision-making processes.
- Look at data analytics service providers and quantify the services' maturity in terms of tools, 7V's of the data, the in-house expertise, and the technical staff skills.
- Analyze the impact of data analysis education, and the influences a data analytics service can contribute to its client's expertise in this subject.

- Quantitative research to measure the impact of the acquired big data analytics capabilities in the company. The arrangement of different capabilities may present different results.
- Dig into factors that support or contribute to the development of big data analytics services. Some possible conditions are the availability of open data and certain companies' contexts, such as existing networks of suppliers and access of data from this network, the products offered, or large access of customer base.
- Explore factors that could infuse data-driven culture in SMEs.

## REFERENCES

- 4property. (2020). 2020 *Property transactions by agencies*. Retrieved from the platform <http://www.ippi.io>.
- Akter S, Wamba SF (2016). *Big data analytics in E-commerce: a systematic review and agenda for future research*. *Electron Mark* 26(2):173–194.
- Akter S, Wamba SF, Gunasekaran A, Dubey R, Childe SJ (2016a) *How to improve firm performance using big data analytics capability and business strategy alignment?* *Int J Prod Econ* 182:113–131.
- Bain, J. S. (1968). *Industrial organization*. New York: Wiley.
- Barney, J. B. (1991). *Firm resources and sustained competitive advantage*. *Economics Meets Sociology in Strategic Management*, 203–227. doi:10.1016/s0742-3322(00)17018-4.
- Barton, D., Court, D., (2012). *Making advanced analytics work for you*. *Harvard Business Review*. 90, 78.
- Beyer, M. A., & Laney, D. (2012). The importance of ‘big data’: a definition. Stamford, CT: Gartner, 2014-2018.
- Bouwman, H., Nikou, S., Molina-Castillo, F. J., & de Reuver, M. (2018). The impact of digitalization on business models. *Digital Policy, Regulation and Governance*.
- EMC Solution Group (2012). *Big data-as-a-service: A market and technology perspective*. Retrieved from: <https://www.yumpu.com/pt/document/read/4544987/big-data-as-a-service-a-market-and-technology-perspective-emc>.
- Carvalho, Luciana Lima Melik de (2019). *CORPORATE INNOVATION THROUGH ENGAGEMENT WITH STARTUPS: SEARCHING FOR SUSTAINED COMPETITIVE ADVANTAGE?* Rio de Janeiro, 2019.
- Corbett, CJ. (2018). *How Sustainable Is Big Data?* *PRODUCTION AND OPERATIONS MANAGEMENT*, 27(9), 1685-1695.
- Christie, B. (1985). *Human Factors of Information Technology in the Office*. John Wiley & Sons, Inc., USA.
- Creswell, John W. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, Calif: Sage Publications, 2003.
- Davenport TH, Harris JG, David W, Jacobson AL (2001) Data to knowledge to results: building an analytic capability. *Calif Manag Rev* 43(2):117–138.

- Demchenko Y, Grosso P, De Laat C, Membrey P (2013) Addressing big data issues in scientific data infrastructure. In: 2013 International conference on collaboration technologies and systems (CTS). IEEE, pp 48–55.
- EMC (2014). *The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things*. Retrieved from <https://www.emc.com/leadership/digital-universe/>.
- Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they?. *Strategic management journal*, 21(10-11), 1105-1121.
- Fosso Wamba, S., Gunasekaran, A., Akter, S., Ren, S. Ji-fan., Dubey, R. & Childe, S. J. (2017). *Big data analytics and firm performance: effects of dynamic capabilities*. *Journal of Business Research*, 70 356-365.
- George G, Osinga CE, Lavie D, Scott B (2016) *Big data and data science methods for management research*. *Academy of Management Journal*, 59(5):1493–1507.
- Ghasemaghaei M, Hassanein K, Turel O (2015) Impacts of big data analytics on organizations: a resource fit perspective. In: *Proceedings of 21st Americas conference on information systems (AMCIS)*.
- Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic management journal*, 17(S2), 109-122.
- Gupta, M., & George, J. F. (2016). *Toward the development of a big data analytics capability*. *Information & Management*, 53(8), 1049-1064.
- Henderson JC, Venkatraman H (1993) *Strategic alignment: leveraging information technology for transforming organizations*. *IBM Syst J* 32(1):472–484.
- Horey, J., Begoli, E., Gunasekaran, R., Lim, S. H., & Nutaro, J. (2012) *Big data platforms as a service: Challenges and approach*. In *Proceedings of the 4th USENIX Conference on Hot Topics in Cloud Computing, HotCloud'12*, pages 16–16, Berkeley, CA, USA, 2012. USENIX Association.
- Kiron D, Shockley R (2011) *Creating business value with analytics*. *MIT Sloan Management Review* 53(1):57–63.
- Kiron D, Prentice PK, Ferguson RB (2014) *The analytics mandate*. *MIT Sloan Management Review* 55(4):1–25.
- Kooper, M. N., Maes, R., & Lindgreen, E. R. (2011). *On the governance of information: Introducing a new concept of governance to support the management of information*. *International Journal of Information Management*, 31(3), 195-200.

- Kumar, N., Stern, L. W., & Anderson, J. C. (1993). *Conducting interorganizational research using key informants*. *Academy of management journal*, 36(6), 1633-1651.
- Kwon O, Lee N, Shin B (2014) Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management* 34(3):387–394.
- Lamba HS, Dubey SK (2015) Analysis of requirements for big data adoption to maximize IT business value in reliability. In: 2015 4th International Conference on Infocom technologies and optimization (ICRITO) (trends and future directions). IEEE, pp 1–6.
- Loebbecke C, Picot A (2015) Reflections on societal and business model transformation arising from digitization and big data analytics: a research agenda. *Journal of Strategic Information Systems* 24(3):149–157.
- Madhani, Pankaj M., (2010). *RESOURCE BASED VIEW: CONCEPTS AND PRACTICES*. ed., pp. 3 -22, Icfai University Press, Hyderabad, India, 2009, Available at SSRN: <https://ssrn.com/abstract=1578704>.
- Maroufkhani, P., Ismail, W. K. W., & Ghobakhloo, M. (2020). *Big data analytics adoption model for small and medium enterprises*. *Journal of Science and Technology Policy Management*.
- Marston, S., Li, Z., Brandyopadyay, S., Zhang, J., Ghalsasi, A., 2011. Cloud computing – the business perspective. *Decis. Support Syst.* 51 (1), 176–189.
- McAfee A, Brynjolfsson E, Davenport TH, Patil DJ, Barton D (2012) *Big data: the Management revolution*. *Harvard Business Review* 90(10):61–67.
- Mell, P., & Grance, T. (2011). *The NIST definition of cloud computing*. Gaithersburg: National Institute of Standards and Technology.
- Mikalef P, Pappas IO, Giannakos MN, Krogstie J, Lekakos G (2016) Big data and strategy: a research framework. In: Proceedings of the 10th mediterranean conference on information systems (MCIS).
- Mikalef, Patrick & Pappas, Ilias & Krogstie, John & Giannakos, Michail. (2017). *Big data analytics capabilities: a systematic literature review and research agenda*. *Information Systems and e-Business Management*.
- Mikalef, Patrick & Pateli, A. (2017). *Information technology-enabled dynamic capabilities and their indirect effect on competitive performance: Findings from PLS-SEM and fsQCA*. *Journal of Business Research*, 70, 1-16.

- Mikalef, Patrick & Boura, Maria & Lekakos, George & Krogstie, John. (2019). *Big data analytics and firm performance: Findings from a mixed-method approach*. Journal of Business Research, 98, 261-276.
- Mikalef, P., Pappas, I. O., Krogstie, J., & Pavlou, P. A. (2020). *Big data and business analytics: A research agenda for realizing business value*. Information & Management, 57(1), 103237.
- Müller O, Junglas I, vom Brocke J, Debortoli S (2016) Utilizing big data analytics for information systems research: challenges, promises and guidelines. European Journal of Information Systems 25(4):289–302.
- O'Brien, C. (2007), *Case Study Research: Principles and Practices by John Gerring*. International Statistical Review, 75: 432-432. doi:10.1111/j.1751-5823.2007.00030\_25.x.
- O'Connor, C. and Kelly, S. (2017). *Facilitating knowledge management through filtered big data: SME competitiveness in an agri-food sector*. Journal of Knowledge Management, Vol. 21 No. 1, pp. 156-179. <https://doi.org/10.1108/JKM-08-2016-0357>.
- OECD (2017a). *Enhancing the contributions of SMEs in a global and digitalized economy*. Meeting of the OECD Council at Ministerial Level.
- OECD (2017b). *Going Digital: Making the Transformation Work for Growth and Well-Being*. Meeting of the Council at Ministerial Level, 7-8 June 2017, retrieved from <https://www.oecd.org/mcm/documents/C-MIN-2017-4%20EN.pdf>.
- OECD (2019). *Strengthening SMEs and Entrepreneurship for Productivity and Inclusive Growth: OECD 2018 Ministerial Conference on SMEs*. OECD Studies on SMEs and Entrepreneurship, OECD Publishing, Paris, Retrieved from <https://doi.org/10.1787/c19b6f97-en>.
- Ogbuokiri, B., Udanor, C. and Agu, M. (2015), "Implementing bigdata analytics for small and medium enterprise (SME) regional growth", IOSR Journal of Computer Engineering, Vol. 17, pp. 35-43.
- Ransbotham, S., Kiron, D., Prentice, P.K., (2015). *Minding the Analytics Gap*. MIT Sloan Management. Review.56,63–68.
- Rasmussen, J. (1987). *Information Processing and Human-Machine Interaction. An Approach to Cognitive Engineering*. North-Holland. North-Holland Series in System Science and Engineering, 12.



- Ravichandran T, Lertwongsatien C (2005) Effect of information systems resources and capabilities on firm performance: a resource-based perspective. *Journal of Management Information Systems* 21(4):237–276
- Rouse, M. (2012). *Big data analytics. Definition from WhatIs.com*. Retrieved from <http://searchbusinessanalytics.techtarget.com/definition/big-data-analytics> [Accessed:24 Jul 2020].
- Schilke O (2014) *On the contingent value of dynamic capabilities for competitive advantage: the nonlinear moderating effect of environmental dynamism*. *Strategic Management Journal* 35(2):179–203.
- Seddon JJ, Currie WL (2017). *A model for unpacking big data analytics in high-frequency trading*. *J Bus Res* 70:300–307.
- Singh, D. (2019). *Implementation of technology innovation in MSMEs in India: case study in select firms from Northern region*. *Journal of Science and Technology Policy Management*, Vol. 10 No. 3, pp. 769-792.
- Talebi, K., Tajeddin, M., Rastgar, A.A. and Emami, A. (2012). *Internationalization of SMEs and organizational factors in a developing country: a case study of ICT industry of Iran*. *International Journal of Academic Research in Business and Social Sciences*, Vol. 7 No. 1, pp. 2222-6990.
- Teece, D., Pisano, G., & Shuen, A. (1997). *Dynamic Capabilities and Strategic Management*. *Strategic Management Journal*, 18(7), 509-533.
- Tien, E.L., Ali, N.M., Miskon, S., Ahmad, N. and Abdullah, N.S. (2020), *Big Data Analytics Adoption Model for Malaysian SMEs*, Springer International Publishing, Cham, pp. 45-53.
- Vajjhala, N. R., & Ramollari, E. (2016). *Big Data using cloud computing – Opportunities for small and medium-sized enterprises*. *European Journal of Economics and Business Studies*, 4, 129–137.
- Venters, W., Whitley, E., 2012. A critical review of cloud computing: researching desires and reality. *J. Inf. Technol.* 27 (3), 179–197.
- Vidgen R, Shaw S, Grant DB (2017) *Management challenges in creating value from business analytics*. *European Journal of Operational Research* 261(2):626–639.
- Zheng, Z., Zhu, J., & Lyu, M. R. (2013, June). *Service-generated big data and big data-as-a-service: an overview*. In 2013 IEEE international congress on Big Data (pp. 403-410). IEEE.

- Wamba SF, Akter S, Edwards A, Chopin G, 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.
- Wamba SF, Gunasekaran A, Akter S, Ren SJF, Dubey R, Childe SJ (2017) *Big data analytics and firm performance: effects of dynamic capabilities*. *Journal of Business Research* 70:356–365.
- Wang, Shouhong & Wang, Hai. (2020). *Big data for small and medium-sized enterprises (SME): a knowledge management model*. *Journal of Knowledge Management*. ahead-of-print. 10.1108/JKM-02-2020-0081.
- Wernerfelt, B. (1984). *A Resource-Based View of the Firm*. *Strategic Management Journal*, 5, 171-180.
- Winter SG (2003). *Understanding dynamic capabilities*. *Strategic Management Journal* 24(10):991–995.
- Wixom BH, Watson HJ, Werner T (2011). *Developing an enterprise business intelligence capability: the Norfolk southern journey*. *MISQ Exec* 10(2):61–71.
- World SME Forum (2017). *Think Big for Small: Small and Medium Enterprises as Pillar for Future-oriented Sustainable Growth*. B20 Cross-Thematic Group, 2017. Retrieved from <http://www.worldsmeforum.org/resources/>.

## APPENDIX

### Appendix A – Interview Script Customers

**Context:** Real estate agencies in Ireland. The majority are small companies (between 1-9 employees). And few cases with medium-sized companies.

**Goal:** Identify the level of these companies' big data analytics capabilities and the impact of the data analytics service on them.

#### 1. Introduction

- a. Would you tell us about what your organization does?
- b. Would you tell us about what you do and your background?

#### 2. Big Data Context

- a. What types of data are held by your organization? Of this data, what would you consider to be "big"?
- b. Who uses the data?
- c. Who administered the data (manages to add/change/delete)?
- d. Do you have big data technologies in your organization besides iPPi? If yes, which one(s)?
- e. Who first introduced these tools to the organization (from within IT, from another function)? (If the answer above was yes)
- f. Who uses these tools? (If the answer above was yes)
- g. Have you adopted any technology for working with big data after using iPPi?

#### 3. Big Data Value Creation

- a. How do people that work with big data create value in the company? How do they influence decision-making? What types of decisions?
- b. How do you see iPPi in the decision-making process of the company? Does the company use data in the report decision-making?

#### 4. Organization

- a. How many people with deep data skills work in your company?
- b. Do you have any formal strategy in the company using big data analytics?

- c. Do you have specific people for dealing with data in your organization? If yes, what is their job role?
- d. Are they generalists or domain specialists?
- e. What types of activities are they involved in?
- f. How do you use this data for doing these analyses? (Try to explore the techniques, the tools, and how is the workflow)
- g. Have you had the same approach before using iPPi (in terms of the tools and the techniques)?

#### **5. Process**

- a. In your organization, do you have practices about data management (retention policies, backup, access to data)?
- b. Do you believe that there is a data-driven culture in your company? What factors have contributed to enabling/hindering this culture? What could be improved?

#### **6. People**

- a. How important do you think analytics skills are for your market (real estate)
- b. Before using iPPi, have your company provided training about analytics topics for its employees.

#### **7. Barriers**

- a. Going forward, what other barriers to using and creating value from big data do you see? E.g., organizational, managerial, technical

#### **8. Value**

- a. What would you say is the value of big data analytics for your company?
- b. Have you been able to capture this value? Is it easy to quantify?
- c. In which areas have big data analytics benefited you most? (e.g., products, services, processes? Or in which way, radical or incremental improvements)
- d. How has the environment in which you compete influenced your decision to adopt data analytics?
- e. Do you think a platform like iPPi can help the organization adopt data analytics? How do you see this process?

## **Appendix B – Interview Script Service Provider**

Context: Service provider of big data analytics service.

**Goal:** Identify the level of the big data analytics capabilities that the service provider can provide.

### **1. Introduction**

- a. Would you tell us about what your organization does?
- b. Would you tell us about what you do and your background?

### **2. Data**

- a. Do your company has access to large, unstructured, or fast-moving data for analysis.
- b. Does your company integrate data from multiple sources into a data warehouse for easy access?
- c. Does your company integrate external data with internal (or client's data) to facilitate business environment analysis?

### **3. Technology**

- a. Has your company explored or adopted parallel computing approaches (e.g., Hadoop) to big data processing?
- b. Do your company has explored or adopted different data visualization tools?
- c. Do your company have explored or adopted new forms of database such as Not Only SQL (NoSQL)
- d. Do your company has explored or adopted cloud-based services for processing data and performing analytics?
- e. Do your company has explored or adopted open-source software for big data analytics?

### **4. Process**

- a. Do you have practices about data management (retention policies, backup, access to data)?
- b. Do you believe that there is a data-driven culture in your company? What factors have contributed to enabling/hindering this culture? What could be improved?

### **5. People**

- a. Our 'Big data analytics' staff has the right skills to accomplish their jobs.
- b. Our 'Big data analytics' staff is well trained.
- c. We provide big data analytics training to our own employees.
- d. Our 'big data analytics' staff has suitable education to fulfill their jobs.