

Data Science Applied to Marketing: A Literature Review

ALBÉRICO ROSÁRIO¹, LUÍS BETTENCOURT MONIZ^{2,4} AND RUI CRUZ^{2,3}

¹GOVCOPP, Escola Superior de Gestão e Tecnologia de Santarém
Instituto Politécnico de Santarém
Santarém, 2001-904 Portugal

E-mail: alberico.rosario@esg.ipsantarem.pt

⁴NOVA Information Management School

Universidade Nova de Lisboa
Lisboa, 1070-312 Portugal

²Universidade Europeia

Lisboa, 1500-210 Portugal

E-mail: luis.moniz@universidadeeuropeia.pt

³Unidcom-IADE

Lisboa, 1200-649 Portugal

E-mail: rui.cruz@universidadeeuropeia.pt

Data Science applied to Marketing has been a research interest due to competitive advantages in business. We have applied a systematic literature review between 2010 and 2020, reaching a total of 19 valid articles. After a deeper segmentation, 13 articles were selected for inclusion in the review comprising the period 2013-2020. On scientific production, the topic Data Science Applied to Marketing, in 2020, has a new subject of interest. The number of citations has been growing since 2015 and the findings revealed that marketing is recurring of a variety of data science methods, from micro-segmentation and real-time application to natural language processing. The impact is evident in digital advertising, micro-segmentation and micro-targeting, speed and performance, and real-time experimentation. The use cases of data analytics in marketing have used four methods with the highest potential to impact marketing approaches: Internet-of-Things, big data, artificial intelligence, and machine learning.

Keywords: data science, marketing, decision making, research trends, literature review

1. INTRODUCTION AND BACKGROUND

The total volume of data consumed and generated online has increased exponentially over the past few years, driven primarily by an increasing ubiquity of consumer digital technologies. There are, presently, more than 6 billion operational Internet-connected consumer devices, which cumulatively generate an excess of a daily 2.5 million terabytes of data [1]. Current estimations also project that there will be approximately 2 megabytes of data generated every second for every single person on Earth by the end of 2020 [10]. These upward trends in information consumption and generation consequently inform the rise to prominence of the discipline of data science.

Big data is often defined by volume, velocity, and variety. To obtain competitive advantages, marketing needs to embed and integrate competencies from several disciplines such as data science, machine learning, text processing, audio-processing, and video-processing. Both big data and data science have impacted the marketing science by developing

Received August 25, 2020; revised November 1 & December 22, 2020; accepted January 19, 2021.
Communicated by Maria José Sousa.

and understating customer choice modeling, customer lifetime value, customer lifetime value, demand forecast, response modeling, brand valuation, consumer behavior and consumer-preference measurement, among others [5].

The marketing benefits from big data are generally accepted and findings show that big data applications in marketing are still in an embryonic stage [1].

Although, currently, there is no consensus regarding a universal definition of the term ‘data science,’ it can be defined as an inter-disciplinary field that uses scientific methods and processes to retrieve knowledge and extract insights from unstructured and structured data [12]. Data science unifies concepts and methods drawn from a variety of scientific disciplines including computer science, mathematics, and statistics. Scholars and practitioners from multiple fields share the widespread recognition that data science offers the mechanisms by which data-driven approaches are to transform every aspect of modern life [10]. The transformational impacts of the ubiquity of data are already being witnessed across multiple industries and sectors including healthcare, supply chain management and logistics [1]. However, no other domain, arguably, has experienced a similar magnitude of impact like the marketing industry. This paper conducts an exploratory review of the insights generated by the latest research concerning the application of data science to marketing and the use of data science as an innovative marketing tool.

2. METHODOLOGY

This paper implements a methodological approach aimed at achieving a comprehensive and systematic review of extant pertinent literature to the exploration of overarching objectives and specific research question. The began with a comprehensive search of literature, consistent with the best practice guidelines concerning the appropriate number of sources of studies during a literature review [3, 13].

The research was carried out in the following configuration: (i) determination of the research question; (ii) location; (iii) selection and evaluation; (iv) analysis and synthesis; (v) results presentation; and (vi) results discussion [13, 14].

As indicated, this paper’s overarching goal is to explore what is the current body of research and literature reveals about the application of data science, as defined above, to the field of marketing. Additionally, the paper seeks to demonstrate the utility of data science as an innovative marketing tool. The subject matter detailed herein is based on an extensive bibliographic research conducted on the SCOPUS database over a month-long duration in June 2020. The choice of SCOPUS database was informed by the need for a robust, yet valid source of information generated by peer-reviewed, empirical, scientific studies. According with Elsevier Scopus is a source-neutral abstract and citation database, curated by independent subject matter experts. It places powerful discovery and analytics tools in the hands of researchers, librarians, institutional research managers, and funders.

The criteria adopted for the literature search on the referred database included, firstly, delimiting the disciplines of interest to business administration and accounting given the close relationship between the two domains with subject matter which is relevant to marketing research. Secondly, and as an opportunity for research, the exclusion of studies published outside the duration between the years 2010 and 2020 was adopted as a criterion. The review process was scoped further by the use of two explicit search terms, namely

“Data Science” and “Marketing,” combined with the Boolean operators “AND” and “OR.”

An initial search of literature generated a total of 23 articles, which upon further subjection to subsequent quality and validity checks was reduced to a total of 19 valid candidate studies. Through further review of the 19 articles’ titles, summaries, abstracts, and keywords, a total of 13 articles were selected for inclusion into the literature review (2013-2020). Thematic analysis was adopted as the most effective approach to synthesize the core subject of the retrieved research since it is a process of identifying patterns or themes within qualitative data. Differently from grounded theory, which sets out to discover or construct theory from data systematically obtained and reviewed using comparative analysis, thematic analysis emphasizes the identification, review analysis and interpretation of meaning patterns, which are more suitable for our study purpose.

Table 1 below provides a summary of the studies and the topics covered by each researcher.

Table 1. Summary of studies included in the analysis of the topics covered.

Theme	Title	Investigators
Application of Machine Learning (ML) to Marketing; and ML-driven marketing tools.	Predicting Customer Quality in E-Commerce Social Networks: A Machine Learning Approach	Ballestar <i>et al.</i> (2019)
ML application and use in marketing.	Constraints-Based Explanations of Classifications.	Deutch <i>et al.</i> (2018)
Application of ML to marketing and ML as a marketing tool.	Model Selection Using Database Characteristics: Developing a Classification Tree for Longitudinal Incidence Data.	Schwartz <i>et al.</i> (2014)
Application of big data (BD) to marketing and BD as a marketing tool.	A Statistical Learning Ontology for Managing Analytics Knowledge.	Behnaz <i>et al.</i> (2019)
Application of BD to marketing and use of BD-powered marketing tools.	Marketing Science and Big Data	Chintagunta <i>et al.</i> (2016)
Application of BD to marketing and BD as a marketing tool.	Introduction	Seshadri (2019)
	Big Data Analytics: A Review on Theoretical Contributions and Tools Used in Literature	Grover <i>et al.</i> (2017)
	Cognitive Computing, Big Data Analytics, and Data Driven Industrial Marketing.	Lytras <i>et al.</i> (2020)
Implications of data science for marketing. BD application to marketing.	Business Challenges and Research Directions of Management Analytics in the Big Data Era	Zhao <i>et al.</i> (2014)
Implications of data science for marketing.	Closing the Skills Gap: Finding Skilled Analytics Professions for A Dynamically Changing Data-Driven Environment	Stanton, & Stanton, (2019)
	Information in Digital, Economic, and Social Networks.	Sundararajan <i>et al.</i> (2013)
Artificial intelligence (AI) and its application to marketing, and AI as a marketing tool.	Artificial Intelligence: Redefining Marketing Management and the Customer Experience	Marinchak <i>et al.</i> (2018)
The Internet of Things (IOT) and its application to marketing, and the IOT as a marketing tool.	The Internet-of-Things: Review and Research Directions	Ng Irene & Wakenshaw (2017)

Source: own elaboration.

3. FINDINGS OF THE THEMATIC ANALYSIS

Graph of Fig. 1 presents publication evolution over the years. Journal with more publications was the Marketing Science with two publications. Others with one publication are: Applied Marketing Analytics; Global Journal of Flexible Systems Management; Industrial Marketing Management; Information Systems Research; International Conference on Information and Knowledge Management; International Journal of e-Entrepreneurship And Innovation; International Journal of Research in Marketing; International Series in Operations Research and Management Science; Journal of Management Analytics; Lecture Notes in Business Information Processing; Review of Managerial Science.

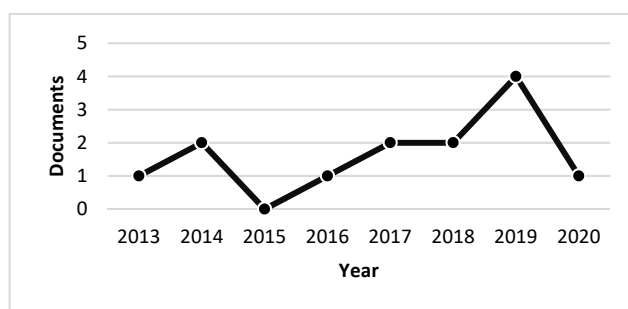


Fig. 1. Documents by year. (Source: own elaboration)

Table 2. Scimago journal and country rank impact factor.

Title	SJR	Best Quartile	H Index
Marketing Science	6,853	Q1	113
Information Systems Research	3,476	Q1	145
International Journal of Research in Marketing	2,910	Q1	95
Industrial Marketing Management	2,080	Q1	125
International Series in Operations Research and Management Science	0,890	Q1	39
Global Journal of Flexible Systems Management	0,786	Q1	21
Review of Managerial Science	0,780	Q1	20
International Conference on Information and Knowledge Management Proceedings	0,610	—*	20
Lecture Notes in Business Information Processing	0,260	Q3	44
International Journal of e-Entrepreneurship and Innovation	0,220	Q3	22
Journal of Management Analytics	0,150	Q3	8
Applied Marketing Analytics	—*	—*	—*

* not available. (Source: own elaboration)

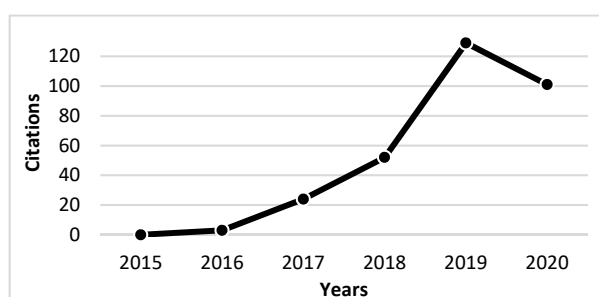
Table 2 refers the Scimago Journal & Country Rank (SJR) the best quartile and the H index by publication, with 7 articles in Q1.

The subject areas covered by the thirteen scientific articles were: Business, Management and Accounting (13); Decision Sciences (6); Computer Science (3); Mathematics (3); Social Sciences (2) and Engineering (1).

References covered by keywords: Data Science (eight references); Machine Learning (four references); Big Data (three references); Business Intelligence; Knowledge Management (two references) and all the following subjects with one reference each:

- Artificial Neural Network; Big Data Technologies; Business Analytics; Business And Societies; Business Models; Cashback; Classification (of Information); Classification Results; Classification Tree; Closing The Skills Gap; Commerce; Commodity Pricing; Computation Theory; Computational Social Scientist; Computer Hardware Description Languages; Computer Science; Consumer Assistant; Consumer Experience; Contagion; Data Analytics; Data Mining; Data Provenance; Data-intensive Computation; Database Constraints; Database Constraints Theory; Decision Trees; Digital Technologies; E-commerce; Forecasting; Generic Approach; Hidden Markov Models; Hierarchical Bayesian Methods; Inference; Information; Information Management; Information Technology; Job Requirements; Learning Systems; Machine Components; Management Techniques; Marketing; Marketing Analytics; Marketing Applications; Marketing Automation; Marketing Campaign; Marketing Intelligence; Marketing Technology; Markets; Model Selection; Modeling; Networks; Networks (circuits); Ontology; Ontology Development; Position Descriptions; Posterior Predictive Model Checking; Predictive Analytics; Predictive Information; Predictive Model; Quantitative Analysis; Random Forest; Reputation Management; Research Traditions; Science Applications; Semantic Modeling; Semantic Technologies; Semantic Technology; Semantics; Service; Social Network; Social Networking (online); Social Scientists; Statistical Learning; Statistics; Supervised Learning By Classification; System Prototype; Systematic Literature Review; User Requirements; and finally Virtual Personal Assistants (VPAs).

The most cited article during this time-period was “*The Internet-of-Things: Review and research directions*” by Ng Irene, & Wakenshaw [11] with 115 citations in the International Journal of Research in Marketing with impact factor 2,910 (SJR), best quartile (Q1) and with H index (95). The article focuses on a review of the Internet of Things (IoT) through four conceptualizations: (i) IoT as a settlement and density of resource information; (ii) IoT as digital materiality; (iii) IoT as an assembly or service system; (iv) and, finally, IoT as modules, transactions and service.



Source: own elaboration

Fig. 2. Evolution of citations between 2015 and 2020.

The h-index was used to measure the productivity and impact of the published work, based on the largest number of articles that had at least the same number of citations. Of the documents considered for the h-index 6, have been cited at least 6 times. In Annex I the citations of all scientific articles from 2013 to 2020 are analyzed, of the thirteen scientific papers three were not cited until 2020 (Table 3).

Table 3. Documents without citations until 2020.

Documents	Investigators	Title	SJR	Best Quartile	H index
Cognitive computing, Big Data Analytics and data driven indu...	Lytras <i>et al.</i> (2020)	Industrial Marketing Management	2,080	Q1	125
Introduction	Seshadri (2019)	International Series in Operations Research and Management Science	0,890	Q1	39
A Statistical Learning Ontology for Managing Analytics Knowl...	Behnaz <i>et al.</i> (2019)	Lecture Notes in Business Information Processing	0,260	Q3	44

Source: own elaboration

In Annex II, we analyzed the self-citation of the documents during the period 2013 to 2020, the seven documents were self-cited 29 times, Big Data Analytics: A Review on Theoretical Contributions and Tools Used in Literature (2017), Global Journal of Flexible Systems Management, was self-cited 16 times in 2017-2; 2018-8; 2019-3; and 2020-3.

3.1 Implications of Data Science for Marketing

The research reviewed herein revealed that ever since the emergence of data science into the mainstream, scholars and practitioners consistently anticipated that marketing would be a focal area of the discipline's impact. Early research suggests multiple aspects marketing science, particularly marketing programs and strategies, that were expected to witness the majority of the practical impact of data science [18, 19]. More recent research validates these earlier predictions and postulations, given that the field of marketing had evolved into the frontier of a variety of data science methods that range from micro-segmentation to natural language processing.

Stanton and Stanton's [17] exploration of the current state of the data-driven business environment reveals that unlike other sectors and industries that have been slower to adopt data science approaches and techniques, the field of marketing has been proactive in its real-time application of data-driven approaches to the large data sets within its possession [17]. According to the authors, the impact of data science in marketing is particularly evident in four types or areas of use cases, namely digital advertising, micro-segmentation and micro-targeting, speed and performance, and real-time experimentation.

Stanton & Stanton [17] observe that the marketing functions of organizations, especially small businesses that were previously locked out of channels of traditional advertising, are increasingly utilizing data-driven approaches to leverage cost-effective marketing mechanisms availed within the digital realm [17]. Data science methods involving statistical analyses of unstructured and structured data are also allowing marketers to bolster the efficiency and effectiveness of their micro-targeting strategies [17]. They enhance the competencies and capabilities relating to how marketers design and deliver specialized offerings to highly specific and smaller customer groups.

Regarding speed and performance, data analytics-led marketing approaches have shown the potential to enhance the execution of marketing campaigns, and there's also tangible evidence suggesting that data science is instrumental to evaluating hence improving the effectiveness of marketing programs [17]. Furthermore, one of the most pronounced impacts of the incorporation of data science into marketing stems from the fact that it has introduced a new set of core competencies for which organizations strive, namely the organization's marketing functions' ability to utilize analytics to better comprehend customer sentiment about products, services, and their attributes [17]. Data science methods introduce this core competency by allowing for real-time testing and experimentation with different customer engagement scenarios, rather than the traditional approaches of intermittent testing or hindsight experimentation.

More specifically, the marketing use cases of data analytics witnessed presently are consistent with previous predictions concerning four methods of data analytics with the highest potential to impact extant marketing approaches, namely Internet-of-Things (IoT), big data (BD), artificial intelligence (AI), and the machine learning (ML) [18, 19]. BD is the field of data science that involves the computation analysis of, and systematic extraction of insights from, extremely voluminous data that are too large and complex to be analyzed by traditional data-processing software and applications [18]. AI, on the other hand, is a domain of data science that deals with the application of machines that programmatically simulate human intelligence [18]. Thirdly, IoT can refer to the component of data science that involves operationalizing all devices connected to automated systems via the Internet [19]. Fourthly, ML in data science involves the application of AI to information systems to accord them with capabilities requisite for automatic access to BD and the generation of meaningful patterns and insights from the said data with minimal to no human intervention [19]. Current use cases of data analytics in marketing focus on acquiring and leveraging capabilities from each of the four domains outlined above, as expounded in the subsequent sections of this thematic analysis.

3.1.1 Marketing and the Internet of Things (IoT)

The studies reviewed herein demonstrate that IoT promises and continues to have a variety of applications and use cases in marketing. Broadly, these use cases can be classified into two categories, namely consumer-facing applications and those related to business processes and infrastructure [2, 7]. Examples of the consumer-facing IoT use cases in marketing include geo-targeting, offerings' personalization, the customization of purchase experiences, tailored pricing, product and offer co-creation, and up selling and cross-selling [9]. Additionally, the application of IoT in marketing encompasses a variety of tasks and operations including understanding consumer behavior, designing and monitoring loyalty program, behavioral and demographic targeting, and dynamic pricing [2]. These IoT marketing applications are in turn facilitated by several technology solutions including Internet-connected constituents (ICCs), solutions for near field communication (NFC), and radio-frequency identification (RFID).

Specifically, this review reveals that the application of data science methods to marketing through the intermediary of IoT tools focuses on enhancing two core areas of contemporary marketing practice, namely research and customer-centrism. Firstly, the augmentation of everyday objects with IoT technologies such as RFID tags and ICCs allow

for the automatic, real-time generation and flow of data not only relating to the operations of the objects onto which they are embedded but also the environments in which the objects operate, and their digital representations [11].

The metadata so-generated present marketers with the opportunity to access and analyze a breadth and diversity of information hitherto unimagined, particularly information concerning customers and their local contexts and situations [11]. Moreover, the depth and detail of the consumer information generated by IOT mechanisms transcend the limitations of the traditional types of data on which marketers previously relied. For instance, ICCs and related sensors can recognize consumers' physical reactions and even emotions as they interact with equipped objects [11]. This increases the visibility to marketers of experiential and consumption contexts, thereby resulting in a substantial enhancement of pertinent research.

A notable benefit of enhanced marketing research is that marketers armed with data generated from IoT-powered sources can enhance their implementation of customer-centric strategies and approaches.

The high visibility of consumer contexts that is ushered by IOT applications increases marketers' insight on how consumers experience the products they purchase and the services they receive [2]. The consensus within the current body of marketing theory and research dealing with the topic of consumer experience is that such experiences are personal, subjective, and multifaceted, and that they also consist of a variety of levels including the rational, physical, spiritual, emotional and sensorial [11]. Information that captures all these complexities and nuances of consumer experience is, therefore, difficult to access with traditional research instruments like customer surveys.

However, IoT data provide a mechanism by which marketers can tap into all these interconnected dimensions by enhancing the visibility of direct consumer experiences during specific interactions with ICCs and sensors [7]. This, in turn, generates new insights into value-in-use and value creation and transforms the relationship between firms and their customers into one of a dynamic of co-creation. Therefore, through its facilitation of the application of IoT to marketing, data science helps to not only identify opportunities for enhancing customer experiences but also model and predict consumer behavior more accurately.

3.1.2 Big data in marketing

As already indicated, big data (BD) is one of the core mechanisms in data science that deals with the computational and systemic analysis of data that are too voluminous and/or complex to be successfully processed via conventional data analysis software and applications. BD is, thus, naturally and inherently linked to the field of marketing, given that the latter is a discipline whose operations depend critically on the ability to process the wide variety of structured and unstructured data about, and generated by, both their organizations and consumers [7].

Three types of big data presently dominate the interest of marketers, namely operational, financial, and customer big data [4]. Customer data includes examples such as names, contact information, purchase and Internet histories, activity on social media and other online communities, and surveys [4]. Secondly, financial data include examples such as the organization's statistics on marketing, sales, costs, and margins [4]. Financial big

data may also include competitors' financial data such as pricing. Thirdly, operational big data are those generated by and associated with organizational processes such as logistics, shipping, and customer relationship management [5] Generally, each type of data is typically sourced from different sources and stored in different locations.

Marketers' increasing attention on the three types of big data mentioned is underpinned by the recognition of the potential BD analytics to significantly enhance: organizations' understanding of their target audiences, performance, and the efficiency of operations [4]. For example, an empirical investigation by Lystras [10] and colleagues shows that BD advertisers to gain more in-depth data on their visitors and hence equips them with superior capabilities in providing more targeted advertising [10].

The opportunities offered by BD in terms of utilizing data-driven approaches to target advertisements more effectively are limitless [10]. For example, marketers can combine data on the climate and on consumer behavior to differentiate the offerings made to different markets based on the different climatic conditions experienced by consumers therein, and so on. Beyond the aforementioned use case in advertisement targeting, the latest research reveals that contemporary BD applications also extend to content marketing. Content marketers are increasingly leveraging BD in bolstering their ability to create more relevant content for their audiences [4]. By gathering and analyzing detailed data on their customers and their respective profiles, creators and publishers can tailor and personalize their content based on specific characteristics and criteria.

Another notable application of BD to real-life marketing scenarios is their application by marketers to the implementation of real-time price adjustments. Although the implementation of effective pricing strategies has always been a top priority for marketers, conventional approaches suffer from inherent limitations related to pricing optimality. BD addresses this limitation by providing the depth and detail necessary to adjust prices based on customer data, including by using normal prices for regular customers and lowering them for price sensitive customers in real time [10]. The majority of current use cases in this regard principally emphasize the superior capabilities of BD to streamline the placing and targeting of coupons, promotions, and other mechanisms of flexible pricing strategies.

In the specific context of financial and operational data related to organizations' marketing functions, potential applications and extant use cases demonstrate that BD promises to introduce superior capabilities in measuring return on investment (ROI) into marketing initiatives and evaluating the performance of marketing programs [5,16]. Notably, BD offers a unified mechanism for consolidating information from all marketing channels and activities and considering their performance concomitantly, including by conducting cost-benefit analyses of all relevant elements without the frequent mistakes that often feature in such comparative analyses [19]. Additionally, because BD analytics are capable of undertaking multifactorial analyses of vast amounts of data, they are capable of not only generating insights about a robust array of variables but also provide for more conclusive results [7]. Broadly, current and potential applications of BD to marketing empirically demonstrate the transformational impact that data science is currently having on the marketing field.

3.1.3 Artificial intelligence in marketing

Presently, efforts to apply AI to marketing and incorporate AI-powered tools into contemporary marketing approaches center on the goal of establishing unified marketing

intelligence strategies. AI enables the achievement of the stated overarching goal by facilitating the integration and harmonization/organization of marketing data, thereby powering faster and more effective decision-making by organizations' marketing functions [9]. Specifically, AI-powered tools have proven, over time, to provide solutions to a core problem frequently faced by most marketers namely, how to connect all of the data available across the marketing ecosystem of tools and partners [9].

AI provides a foundational layer upon which to build a holistic view of all customer touchpoints [17]. Most application programming interfaces (APIs) that exist as alternative solutions for marketers to manage their engagement and campaigns data are unique in design hence disparate in functionality [17]. This high degree in software diversity implies, in turn, that marketers quickly encounter limitations to the exclusive use of APIs in connecting marketing data together regardless of the combinations and permutations of APIs that they select [17]. The limitations to connecting data across the marketing ecosystem are also rendered more pronounced by the increasing number of available marketing platforms and sources of marketing data without APIs.

AI addresses the API limitations above directly by offering tools and mechanisms by which any data needed can be incorporated into any given marketing data ecosystem, simultaneously and instantaneously [17]. AI-powered data connection tools act like the human eye, recognizing the sources of relevant data and automatically identifying the ways in which different sources of information should be mapped into extant data models [9]. Without this critical function provisioned by AI, the data connections within marketing data ecosystems are prone to breaking, thereby translating to more inefficiency in the form of wasted time and additional resource requirements for fixing and maintenance.

Therefore, the application of AI to the integration of marketing data can effectively replace the previously inefficient multiplicity of systems, one-off dashboards, and spreadsheets; while at the same time introducing numerous benefits to the marketing function, including the seamless introduction of new data sources into marketing ecosystems and the switching between data platforms in the marketers' stack without the hitches commonly experienced previously.

Additionally, beyond the integration and connection of marketing data, getting to a unified marketing intelligence strategy also requires comprehensive data harmonization through an effective data model. Such harmonization through data models ensures that every piece of interconnected insight on investment, performance, and outcomes is accessible, observable, and retrievable at any given instant [17]. Prior to AI, generating insights from the marketing data ecosystem required the creation from scratch of data models capable of answering specific questions of interest [9]. This would consequently translate to vast time and human capital requirements that would in turn imply continuous investments by organizations.

Given the limitations inherent to the data analysis paradigm introduced above, most organizations today have shifted to the alternative of using dynamic data models that expand their respective bodies of knowledge continually from across customers [17]. This shift has, accordingly, allowed for the automatic mapping into these dynamic models of AI-integrated data, thereby eliminating the need to construct analytical models from scratch [19]. Therefore, AI-assisted marketing allows for not only the integration of data across marketing ecosystems, but also for the data's centralization and for the continuous agility of dynamic models as marketing data continually scale and change.

3.1.4 Machine learning in marketing

Although often used interchangeably, the key distinction between machine learning and AI as data science methods is that ML focuses on the analysis and optimization of processes and problems, rather than the simulation of human intelligence. In this regard, ML has numerous use cases and potential applications in the field of marketing. The literature reviewed herein shows that many organizations the world over are increasingly relying on ML to improve the level of degree of personalization and customization of their marketing approaches [2].

This application of ML to marketing has accelerated the accuracy, precision, and context of optimized message targeting in addition to facilitating real-time personalized advertising across digital platforms [6]. The application of ML to marketing personalization and customization has also resulted in a significant uptake of Sales Qualified Lead generation by marketers, who view ML as being increasingly instrumental to the objectives of increasing win rates and reducing sales cycles [2].

ML also offers marketers operating in high-churn industries like telecommunication with a data-driven tool for analyzing customer churn and streamlining risk and intervention models, thereby significantly reducing the churn rate [15]. By facilitating the construction of evidence-based risk prediction and intervention models, ML allows marketers to consider, for example, the amount of customer lifetime value at play and how an intervention under consideration could potentially affect the probability of churn.

Another notable application of ML to marketing is its utility in the creation and fine-tuning of propensity models that underpin up-sell and cross-sell strategies. ML facilitates the creation and fine-tuning of propensity models by customer segment, customer persona, and product line among other criteria [2]. For example, the definition of products and services with the highest likelihood of being purchased through the construction and application of propensity models is a common practice among today's marketers [2].

ML provides a way of bolstering the previous types of manual propensity models through automation. As in the case of their propensity counterparts, marketers are also increasingly using ML to enhance the ability of predictive models to generate more ideal customer profiles and more accurate lead scoring regimes [6]. In addition to improving lead scoring accuracy, ML-assisted marketing also enhances the optimization of the marketing mix.

4. CONCLUSION

This literature review conducted and restricted to publications referenced in the SCOPUS database during the period 2013-2020, has explored marketing applications and tools across four domains of data science, namely *big data*, *artificial intelligence*, *machine learning* and *Internet of Things*.

It finds that IoT and related data analytics currently find application in a variety of ways aimed at increasing the depth and effectiveness of marketing research, while also enhancing customer experiences' visibility to marketers. This reality has allowed marketing professionals to be more vigilant about consumer behavior and experience in real time, enabling a greater adaptation of the product or service to the characteristics of each con-

sumer, or promoting faster product development or incremental innovation cycles. At the same time, IoT also provides strong possibilities for competitive cooperation between organizations in the same field or industry through the possible connectivity of their value proposals in a complementary way contributing, through knowledge sharing, to the expansion and balance of markets and the reduction of competitive erosion between companies.

Secondly, the review established that the data science domain of BD presents marketers with a wide array of tools that allow for applications such as targeted advertising, content customization, and real-time pricing adjustments. Permanent access to information collected instantly can contribute to the development of different forms of just-in-time associated with product or service development. It is this same access and above all the speed in this access that has made all of us beta testers of software that is widely spread in its most basic version and systematically improved by successive, and in many cases, daily updates that are nothing more than the fruit of information shared by consumers. And if in the beginning informatics has applied this principle, the reality of its potential can be extended to the automobile industry, home appliances or home automation just to mention some areas where its potential can be explored. Also, in the service sector, the possibilities of BD can be very important: airlines, cruise companies, hotel and tourism groups, restaurants, can truly offer tailor-made trips, tailor-made stays, tailor-made meals from the immediate management of information both for previous and new customers.

Thirdly, AI was found to provide tools and applications that facilitate the achievement of unified marketing intelligence strategies through the integration and harmonization of data across the marketing ecosystem. This enormous information management capacity may eventually be extended to the domain of financial and risk management, which will enable large recipients of information such as Google, Amazon or AliBaba, on a planetary scale, or telecommunications operators, local internet providers or even regional distributors and wholesalers, to develop systems that will open them the doors to working in credit services for large numbers of middle class consumers or small and medium enterprises, doing to banks in the financial sector and Uber has done to the taxi industry.

Finally, marketing use cases of ML include tools and applications for personalizing marketing strategies, managing customer churn rates, and optimizing propensity and predictive marketing models. But as learning is accentuated through increasingly powerful algorithms and neural networks as an integral part of ML, greater perspectives are opened with potential impact on customer service and after-sales service, making it possible to quickly identify the type and origin of complaints and the possibilities of conflict, and typify customer segments and typologies with different types of potential and behavior.

To conclude, extending the present analysis to a wider field in terms of Web of Science may remedy some possible weakness in the limitation of the study which, despite the importance and relevance of the Scopus database, may suffer from some conditioning in terms of literature. Extending the research to WoS may therefore be a line of research to be pursued. In the same way, replicating this analysis in a segmented way by sectors, namely in services, with emphasis on tourism and hotels, restaurants and aviation may allow us to identify trends and establish interesting sectoral comparisons.

Annex I: Overview of document citations period 2013 to 2020.

Documents		2013	2014	2015	2016	2017	2018	2019	2020	Total
Cognitive computing, Big Data Analytics and data driven indu...	2020	–	–	–	–	–	–	–	–	–
Predicting customer quality in e-commerce social networks: a...	2019	–	–	–	–	–	–	4	1	5
Closing the skills gap: Finding skilled analytics profession...	2019	–	–	–	–	–	–	–	1	1
Introduction	2019	–	–	–	–	–	–	–	–	–
A Statistical Learning Ontology for Managing Analytics Knowl...	2019	–	–	–	–	–	–	–	–	–
CEC: Constraints based explanation for classifications	2018	–	–	–	–	–	–	1	–	1
Artificial intelligence: Redefining marketing management and...	2018	–	–	–	–	–	–	–	2	2
Big Data Analytics: A Review on Theoretical Contributions an...	2017	–	–	–	–	3	12	18	14	47
The Internet-of-Things: Review and research directions	2017	–	–	–	–	5	26	49	35	115
Marketing science and big data	2016	–	–	–	–	2	–	10	6	18
Business challenges and research directions of management an...	2014	–	–	3	10	4	6	6	3	32
Model selection using database characteristics: Developing a...	2014	–	–	–	4	5	4	1	–	14
Information in digital, economic, and social networks	2013	4	4	8	10	13	8	6	4	57
Source: own elaboration	Total	4	4	11	24	32	56	95	66	292

Annex II: Overview of document self-citation period 2013 to 2020.

Documents		2013	2014	2015	2016	2017	2018	2019	2020	Total
Cognitive computing, Big Data Analytics and data driven indu...	2020	–	–	–	–	–	–	–	–	–
Predicting customer quality in e-commerce social networks: a...	2019	–	–	–	–	–	–	1	–	1
Closing the skills gap: Finding skilled analytics profession...	2019	–	–	–	–	–	–	–	–	–
Introduction	2019	–	–	–	–	–	–	–	–	–
A Statistical Learning Ontology for Managing Analytics Knowl...	2019	–	–	–	–	–	–	–	–	–
CEC: Constraints based explanation for classifications	2018	–	–	–	–	–	–	–	–	–
Artificial intelligence: Redefining marketing management and...	2018	–	–	–	–	–	–	–	–	–
Big Data Analytics: A Review on Theoretical Contributions an...	2017	–	–	–	–	2	8	3	3	16
The Internet-of-Things: Review and research directions	2017	–	–	–	–	–	2	–	1	3
Marketing science and big data	2016	–	–	–	–	1	–	–	–	1
Business challenges and research directions of management an...	2014	–	–	–	2	–	–	–	–	2
Model selection using database characteristics: Developing a...	2014	–	–	–	2	2	–	–	–	4
Information in digital, economic, and social networks	2013	–	–	–	1	1	–	–	–	2
Source: own elaboration	Total	–	–	–	5	6	10	4	4	29

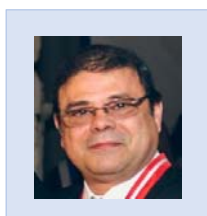
ACKNOWLEDGEMENTS

We would like to express our gratitude to the Editor and the Referees. They offered extremely valuable suggestions or improvements. The authors were supported by the GOVCOPP Research Unit of Universidade de Aveiro and UNIDCOM/IADE research unit of Universidade Europeia.

REFERENCES

1. A. Amado, P. Cortez, P. Rita, and S. Moro, "Research trends on big data in marketing: A text mining and topic modeling based literature analysis," *European Research on Management and Business Economics*, Vol. 24, 2018, pp. 1-7.
2. M. T. Ballestar, P. Grau-Carles, and J. Sainz, "Predicting customer quality in e-commerce social networks: a machine learning approach," *Review of Managerial Science*, Vol. 13, 2019, pp. 589-603.
3. W. Bandara, S. Miskon, and E. Fielt, "A systematic, tool-supported method for conducting literature reviews in information systems," in *Proceedings of the 19th European Conference on Information Systems*, 2011, pp. 1-13.
4. A. Behnaz, M. Bandara, M. Rabhi, and M. Peat, "A statistical learning ontology for managing analytics knowledge lecture notes," *Business Information Processing*, Vol. 345, 2019, pp. 180-194.
5. P. Chintagunta, D. M. Hanssens, and J. R. Hauser, "Marketing science and big data," *Marketing Science*, Vol. 35, 2016, pp. 341-342.
6. D. Deutch and N. Frost, "CEC: Constraints based explanation for classifications," in *Proceedings of International Conference on Information and Knowledge Management*, 2016, pp. 1879-1882.
7. P. Grover and A. K. Kar, "Big data analytics: A review on theoretical contributions and tools used in literature," *Global Journal of Flexible Systems Management*, Vol. 18, 2017, pp. 203-229.
8. M. Lytras, A. Visvizi, X. Zhang, and N. R. Aljohani, "Cognitive computing, big data analytics and data driven industrial marketing," *Industrial Marketing Management*, Vol. 90, 2020, pp. 663-666.
9. C. M. Marinchak, E. Forrest, and B. Hoanca, "Artificial intelligence: Redefining marketing management and the customer experience," *International Journal of E-Entrepreneurship and Innovation*, Vol. 8, 2018, pp. 14-24.
10. P. Moscato and N. J. de Vries, "Marketing meets data science: Bridging the gap," *Business and Consumer Analytics: New Ideas*, Springer, Cham, 2019, pp. 3-117.
11. I. C. L. Ng and S. Y. L. Wakenshaw, "The Internet-of-Things: Review and research directions," *International Journal of Research in Marketing*, Vol. 34, 2017, pp. 3-21.
12. F. Provost and T. Fawcett, *Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking*, O'Reilly Media, Inc., Newton, MA, 2013.
13. A. Rosário and R. Cruz, "Determinants of innovation in digital marketing, innovation policy and trends in the digital age," *Journal of Reviews on Global Economics*, Vol. 8, 2019, pp. 1722-1731.

14. A. Sacavém, R. Cruz, M. Sousa, and A. Rosário, “An integrative literature review on leadership models for innovative organizations, innovation policy and trends in the digital age,” *Journal of Reviews on Global Economics*, Vol. 8, 2019, pp. 1741-1751.
15. E. M. Schwartz, E. T. Bradlow, and P. S. Fader, “Model selection using database characteristics: Developing a classification tree for longitudinal incidence data,” *Marketing Science*, Vol. 33, 2014, pp. 188-205.
16. S. Seshadri, “Introduction,” *International Series in Operations Research and Management Science*, Vol. 264, 2019, pp. 1-15.
17. A. D. Stanton and W. W. Stanton, “Closing the skills gap: Finding skilled analytics professionals for a dynamically changing data-driven environment,” *Applied Marketing Analytics*, Vol. 5, 2019, pp. 170-184.
18. J. L. Zhao, S. Fan, and D. Hu, “Business challenges and research directions of management analytics in the big data era,” *Journal of Management Analytics*, Vol. 1, 2014, pp. 169-174.



Albérico Travassos Rosário, Ph.D. Marketing and Strategy of the Universities of Aveiro (UA), Minho (UM) and Beira Interior (UBI). With affiliation to the GOVCOPP Research Center of the Universidade of Aveiro. Master in Marketing and Bachelor's degree in Marketing his research interests are centred on marketing, innovation and entrepreneurship.



Luíz Bettencourt Moniz, M.Sc. Marketing Analytics and Metrics. Master in Information Management on Marketing Intelligence from NOVA Information Management School (IMS), Lisbon, and Graduate in Statistics and Information Management from NOVA IMS his research interests focus on digital marketing, marketing performance and marketing metrics.



Rui Manuel Nunes Cruz, Ph.D. in Management and Marketing from the Universidad de Sevilla, Spain, also has a DEA from the same university and an MBA in Strategic Foresight at IESFF, Lisbon. Integrated Researcher at UNIDCOM/IADE – Universidade Europeia, Lisbon, his research interests are related with marketing, innovation and knowledge management.