

Topic Modeling Emerging Trends for Business Intelligence in Marketing: With Text Mining and Latent Dirichlet Allocation

Rouhollah Bagheri^{1*}, Nahid Entezarian²

Abstract

This paper examines recent literature in the quest to uncover emerging patterns in the use of business intelligence in marketing. We conducted searches in pertinent academic journals and identified 1044 articles published between 2000 and 2023. To sift through this substantial body of work, we employed text mining techniques to extract pertinent terms in the realms of business intelligence and marketing. Additionally, we applied latent Dirichlet allocation modeling to categorize the articles into various pertinent topics. This analysis was performed within the domains of marketing and business intelligence. This approach enabled us to discover connections between terms and topics, which in turn allowed us to generate hypotheses regarding future research directions. To validate these hypotheses, we gathered and closely examined relevant articles. By pinpointing current research areas, this study underscores potential avenues for future investigation. The findings reveal that the predominant trend in business intelligence applications for marketing is the utilization of business intelligence systems, with a particular emphasis on marketing planning to enhance marketing strategies. Additionally, there is considerable interest in areas such as pricing models for marketing, enhancing brand value through effective social media marketing, employing predictive algorithms for customer data analysis, and harnessing big data for marketing analytics.

Keywords: *Marketing, Business intelligence, Text mining, Latent Dirichlet allocation, topic modeling*

Introduction

Business intelligence (BI) and market information (MI) have become essential tools in the business world. A vast amount of data is generated daily to extract insights from various sources, understand the strategic direction, and formulate successful strategies (Shao et al., 2022).

In the fiercely competitive global business landscape, organizations must constantly pursue accurate information to make well-informed decisions and uphold their competitive advantage over their competitors (Chen & Lin, 2021). Hence, every business entity must analyze and forecast market behavior in order to maintain stability in the midst of market fluctuations and their

capacity to adapt. To achieve this, they must modernize their business processes by incorporating intelligent business approaches and leveraging cutting-edge technologies like business intelligence and data mining. Definitive insights derived from business intelligence can serve as a foundation for significant changes and the adoption of pivotal decisions, including establishing new collaboration opportunities, acquiring new customers, identifying fresh markets, and introducing new products to their customer base.

Business intelligence for sales performance equips sales managers and their teams with valuable analytical data concerning products, pricing trends,

1*. Department of Management, Faculty of Economics and Administrative Sciences, Ferdowsi University of Mashhad, Mashhad, Iran (Corresponding Author: Rbagheri@um.ac.ir)

2. Department of Management, Faculty of Economics and Administrative Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

customer demographics, regions, and sales activities. This data includes information on sales volumes during various time periods, enabling timely decision-making to maintain product availability and sales levels, thus preventing revenue loss. It not only reduces sales attrition but also sustains or increases the organization's profitability. Business intelligence integrated within the sales system is also crucial for assessing how effectively an organization is progressing toward its predefined objectives (Huang et al., 2022). To adapt to changing market dynamics, organizations should periodically scrutinize shifts in customer purchasing behavior. A comprehensive business intelligence system for marketing operations requires precise and relevant market information. However, such data is frequently dispersed across different systems, making access challenging. Consequently, acquiring the necessary information often involves analyzing data from diverse sources.

While there have been surveys and reviews in the marketing field, there is a noticeable dearth of research on the applications and utilization of business intelligence in marketing. In fact, very few studies of this nature are available. As far as we are aware, there has been no direct focus on business intelligence applications in marketing within articles (Bahrami et al., 2012). Additionally, none of the reviews discussed have utilized automated text analysis techniques such as Text Mining (TM), as demonstrated in this study. Text Mining can aid in the analysis of a much wider array of sources.

This study aims to achieve the following objectives: (i) emphasize the key research on the intersection of business intelligence and marketing; (ii) illustrate the evolution of these studies over time; and (iii) discuss the primary findings. To accomplish these goals, we employed a text-mining approach to analyze 1044 indexed articles from Scopus.

The paper conducts an automated text mining literature analysis spanning from 2000 to 2023, focusing on the application of business intelligence in the marketing domain. The first significant contribution of

this study is the identification and presentation of the five most pertinent topics that categorize discussions on the convergence of business intelligence and marketing, providing a structured overview of existing literature, and revealing key authors and prominent journals. The second contribution lies in the classification of topics within the business intelligence and marketing field, the distribution characteristics of these topics across various publication sources and regions, and how these topic distributions change over time (Chen & Lin, 2021). The third contribution provides an overview of the existing literature on business intelligence applications in marketing. The fourth contribution offers a research agenda for future studies exploring the utilization of business intelligence in the realm of marketing. These contents are valuable for newcomers to the field of business intelligence in marketing, providing a comprehensive understanding, and offer insights to researchers.

This study is structured into five main sections. The second section introduces the fundamental concepts related to marketing and business intelligence and references other relevant literature analyses. The third section outlines the methodology used for data collection and analysis. Following that, we present the results of our data analysis in the subsequent section. The final section encompasses an overall discussion, conclusions, and implications regarding the utilization of business intelligence in marketing.

Theoretical background

In the contemporary and highly competitive business landscape, it is crucial for companies to adjust and endure amidst the swift shifts both within and outside their organizations. With the world evolving into a more dynamic and cooperative space, it is essential to construct a sound business strategy. This strategy should aim to provide practical insights and offer guidance to top-level executives, managers, and employees for timely problem-solving and planning

(Sahani, 2018). The notions of Business Intelligence and Marketing play significant roles in enhancing the viability and competitiveness of businesses. They contribute to bolstering sales and improving the efficacy of decision-making within the realm of business operations. This section elaborates on both of these vital concepts.

Business Intelligence

Business Intelligence is the process of extracting, transforming, managing, and analyzing vast datasets through the creation of mathematical models to acquire information and knowledge essential for making decisions in complex scenarios. Business Intelligence encompasses a range of concepts and techniques utilized by organizations to enhance decision-making by employing fact-based support systems. Business Intelligence System research integrates numerous topics, including Data Warehousing, Decision Support Systems, business analysis tools, and content or knowledge management. These systems heavily depend on online analytical processing (OLAP) and data mining methods.

In contrast, business analytics refers to the technology and information systems that enable the analysis and reporting of Big Data using various analytic methods. Business Intelligence has diverse applications, such as determining product pricing, optimizing product placement, and forecasting sales. Forecasting is a discipline that predicts future conditions by extrapolating past data into the future using both subjective estimates and mathematical models.

Business Intelligence can help companies enhance their competitive advantage by effectively leveraging data and converting it into actionable insights. BI solutions provide authorized users with access to data and empower them to correlate it with competitive intelligence within a secure, centrally managed data warehouse (Tilak, 2020; Ghouschi, 2021)

Over recent years, Business Intelligence has garnered significant attention in academia, e-commerce, and the business

world. It has become a crucial tool for improving enterprise performance and driving the growth of e-services and e-commerce. However, due to the rapid advancements in big data and data analytics, BI is facing new opportunities and challenges. Utilizing big data analysis to further enhance BI has become a critical concern for businesses, e-commerce, organizations, and information systems.

The application of the Business Intelligence concept offers substantial opportunities for companies (Chen & Lin, 2021) to increase profitability and reduce risk. BI is recognized as a powerful tool for enhancing operational capabilities. It helps collect vital information from diverse unstructured data sources and transforms them into actionable insights that enable organizations to make informed policy decisions, enhance business efficiency, and boost productivity. The intense competition in the business world underscores the crucial importance of data analysis in the decision-making process. Business intelligence and communication technologies play a vital role in helping organizations achieve organizational sensing agility, decision-making agility, and acting agility in various organizational and environmental contexts. Effective data analysis supports organizations in shaping policies, executing strategic actions, and making business-related decisions (Kim, 2014). As organizations continue to seek ways to improve business performance management, the adoption of business intelligence becomes increasingly widespread. Additionally, organizations adopting big data analytics are shown to make decisions more swiftly and perform better financially compared to their competitors, thanks to the insights gained from business intelligence.

Business Intelligence is a strategic component of organizational design that revolves around the rapid analysis of information to facilitate precise and informed decision-making in a condensed timeframe. This approach involves employing a suite of analytical and functional applications. Through a BI system, an organization can

swiftly identify real-world challenges and their interconnections within its business operations, subsequently prompting the organization to take practical actions aimed at achieving its objectives (Haseli et al., 2020).

Business intelligence tools offer organizations a wide array of applications, including the capacity to explore new markets, assess the demand for their products and services across different market segments, and appraise the effectiveness of their marketing efforts.

Marketing

According to the definition presented by the American Marketing Association in 2007, marketing encompasses the activities, institutions, and processes that revolve around the creation, communication, delivery, and exchange of products or services, all with the goal of delivering value to customers, clients, partners, and society at large.

In the contemporary business landscape, several significant macro-environmental factors, such as digitization, sustainability, and the complex interplay between globalization and anti-globalization trends (as outlined by Bremmer (2014)), are introducing a high degree of uncertainty and turbulence for organizations. Three key shifts in marketing practices and principles are driving a compelling argument for reevaluating the core purpose, assumptions, and implicit models that have traditionally defined marketing, as proposed by Webster and Lusch in (2013.): (i) There's a growing recognition of the impact of marketing at the level of entire service ecosystems, moving beyond the traditional focus on the interaction between the company and the individual customer. (ii) The nature of market offerings is evolving from tangible goods to services and customer experiences, with an increased emphasis on customer involvement in value creation. (iii) Marketing is increasingly being seen as a two-way street, with a focus on its impact on society and how society, in turn, influences marketing. This shift is promoting a move

from a market-oriented approach to a sustainable market orientation. (iv) The role of marketing within organizations is also undergoing transformation (Manian & Ronaghi, 2015). To navigate these dynamic changes, forward-thinking companies in the business world diligently track market developments and align their strategies with evolving market dynamics. This frequently entails the creation and implementation of strategic marketing systems.

Marketing systems that are in perfect alignment with top-tier companies offer market opportunities that lead to more effective market administration and management. This comprehensive market management process comprises four primary steps: (i) Market Analysis: The market environment is intricate and ever-changing, bringing forth both opportunities and new threats. Companies and their strategic systems must continually monitor this dynamic environment, necessitating a wealth of information, including consumer behavior and purchasing patterns. Information systems and market research are used to assess and evaluate consumer and institutional markets. (ii) Selection of Target Markets: Given the diversity in consumer preferences and the inability of any one company to satisfy all consumers, it's crucial for companies to maximize their potential and choose the most suitable market position (Nezhad et al., 2023; Tilak, 2020). This entails a thorough examination of four key aspects: demand analysis, market share projection, target market selection, and market positioning. (iii) Integration of Marketing Elements: This step revolves around the fundamental concept of the marketing mix, which involves managing a set of controllable variables that interact in the market to elicit a desired response. These variables include product design, distribution, pricing, and advertising, all of which play a vital role in breaking into new markets. (iv) Monitoring the Market and Competitors' Activities: This stage involves scrutinizing competitors, their market strategies, and the planning, execution, and control of marketing programs. A crucial component of this

process is to continuously monitor and analyze competitors' actions, product prices, distribution methods, and advertising tactics to stay abreast of the competitive landscape and prevailing market conditions.

In the realm of marketing intelligence, which focuses on the marketing-related facets of business intelligence, information pertaining to a company's markets is gathered and processed to produce insights that facilitate decision-making. Traditionally, marketing intelligence has depended on market surveys as a means to comprehend consumer behavior and enhance product design. For instance, companies utilize consumer satisfaction surveys to investigate customer feelings. With the help of advanced big data analytics technology, critical elements for making strategic marketing decisions, such as understanding customer opinions about a product, service, or company, can now be automatically monitored by extracting insights from social media data (Kim, 2014; Nazarian, 2023).

However, despite the vast potential that big data accessibility offers in the realm of marketing intelligence, it also brings challenges for both practitioners and researchers. Big data analytics primarily wrestles with three major categories of challenges: storage, management, and processing (Peres et al., 2023). When it comes to conventional marketing intelligence activities like extracting customer opinions, modern companies have numerous options for gathering data from various sources, including social media data, transactional data, survey data, sensor network data, and more. Depending on the characteristics of the collected data, different methods can be employed to uncover marketing intelligence. Analytical models developed based on a single data source may offer only limited insights, potentially leading to biased business decisions (Saini, 2022). On the other hand, combining varied information from multiple sources offers a holistic view of the domain and results in more precise

marketing intelligence. Regrettably, harmonizing substantial data from numerous sources to generate marketing insights is a challenging endeavor. This calls for the exploration of innovative techniques, applications, and frameworks to effectively manage big data within the context of marketing intelligence.

Literature analysis

A review of the literature, as outlined by Webster and Lusch in (2013), allows for the examination of a specific topic and the identification of research trends and potential gaps, which can stimulate new studies and discoveries. This step is considered crucial and serves as a foundation for uncovering fresh insights into a research subject, ultimately driving the generation of novel findings. This significance is evident in the numerous publications on the process of conducting literature reviews across various scientific disciplines, as highlighted by (Bahrami et al., 2012).

Traditionally, comprehensive literature analyses have required significant effort from researchers in their quest to understand the current state of knowledge on a particular subject, thereby serving as a catalyst for further research. Thanks to new technologies, online library databases are now easily accessible from any location, offering researchers access to a vast array of available articles. These databases offer search queries that assist in finding articles related to a specific topic. Nevertheless, the sheer number of articles retrieved can pose the daunting challenge of reading the complete content of each paper. In this regard, it is worth noting that smaller sections of articles, such as titles, abstracts, and keywords, can provide initial insights into the research conducted. To tackle this challenge, several text mining (TM) methods have been proposed in recent times for analyzing literature. Table 1 provides a summary of four frameworks for literature analysis that employ various techniques.

Table 1.
Analysis literature of business intelligence in marketing

Subject	Method	Search	Reference
Business intelligence, Marketing information systems	literature review	Not specified	(Stone & Woodcock, 2014)
business intelligence, marketing	Q method	Not specified	(Kim, 2014)
Business intelligence (BI), Digital marketing	Critical review	Not specified	(Tilak, 2020)
business intelligence, marketing	descriptive-survey research and field study	Not specified	(Huang, 2022; Skálén, 2023)
business intelligence; marketing intelligence	Critical review	Not specified	(Webster & Lusch, 2013)

Two articles within the list provide critical evaluations, while the other two articles present literature reviews. One of them utilizes the Q method, and the other employs descriptive-survey research and field study methods. Notably, none of the articles employed text mining to evaluate the literature.

Methodology

The objective of this section is to introduce the research methodology framework, which is visually represented in Figure 1. This section is organized into three components: (1) Text Mining and Topic Modeling, (2) Data Gathering and Preprocessing, and (3) Topic Modeling using Latent Dirichlet Allocation (LDA).

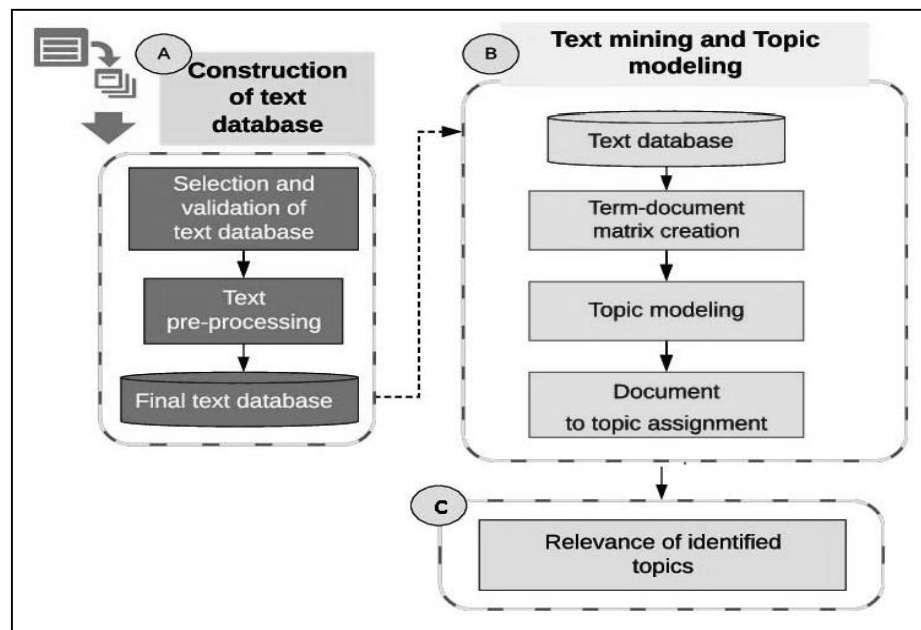


Figure 1. *Methodology of Text Mining and Topic modeling (Buenano-Fernandez et al., 2020)*

Text Mining and Topic Modeling

Text mining is a specialized data mining approach with the goal of extracting valuable insights, including relationships, patterns, and trends, from unstructured or semi-structured data, particularly text documents. The primary procedure in text mining involves converting text into numerical data

through statistical methods to create an organized document-term matrix. This matrix comprises two key dimensions: the words (or terms, consisting of n words) and the documents (Moro et al., 2019).

In this study, the researchers employ a statistical topic model known as Latent Dirichlet Allocation (LDA) to identify latent

topics within the literature review. LDA has been widely utilized to uncover hidden themes in scholarly literature documents.

In addition to LDA, there have been various techniques developed in recent years to generate knowledge through text mining. One such approach is Latent Semantic Analysis (LSA), which leverages natural language processing to identify relationships between textual terms and documents by assuming that words with similar meanings appear in similar text passages. Nevertheless, comprehending the outcomes can be demanding. Taking inspiration from Latent Semantic Analysis (LSA), Probabilistic Latent Semantic Analysis (PLSA) was introduced with the goal of conducting topic analysis on document collections through the application of probabilistic generative models (Barrera & Shah, 2023). Topic modeling is the process of transforming relevant words and their frequencies into an organized structure, where documents are categorized into several topics. While both Probabilistic Latent Semantic Analysis (PLSA) and Latent Dirichlet Allocation (LDA) are probabilistic models, LDA takes into account prior probability distributions during text generation to prevent overfitting of results. The Correlated Topic Model (CTM) by (Blei & Lafferty, 2007) considers the correlations between topics, addressing limitations in

previous models that only focus on probability distributions. However, it is less sensitive to the number of topics and may generate an excessive number of topics, reducing the interpretability and breadth of identified themes. Additionally, CTM requires substantial computing power, making it unsuitable for this study. The Dynamic Topic Model (DTM) introduced by (Blei & Lafferty, 2006) captures topics that change over time, assuming dynamic shifts in topic content. However, the current research concentrates on identifying broader trends and distributional characteristics of research topics, without delving into internal structural changes. Consequently, DTM is not chosen for this study.

It is worth noting that estimating the parameters of the LDA model, a complex optimization problem, typically relies on approximate solutions (Blei et al., 2003), suggesting some trade-offs in the modeling process. The authors emphasize that LDA is the preferred tool for handling multiple topics, as it can determine the probability of each document belonging to different topics and group documents into the most likely matching topics (Blei et al., 2003). Figure 2 illustrates the generative process used by LDA, a widely adopted method for uncovering latent topics in textual materials in the current study.

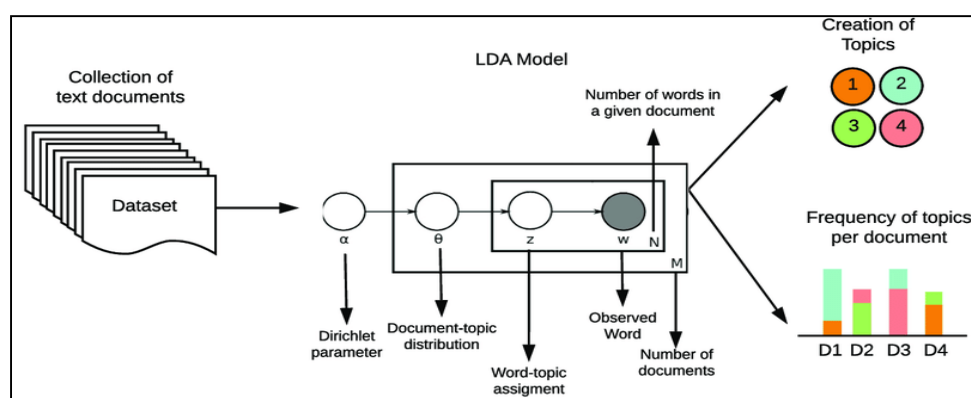


Figure 2. Schematic of LDA algorithm (Buenano-Fernandez et al., 2020)

Data Collection (Selecting articles) and preprocessing

The majority of scholarly academic literature is typically found in peer-reviewed

journals, and the most pertinent ones are included in the Scopus database, as noted by Moro et al. in (Moro et al, 2015). Due to its extensive coverage and larger database,

Scopus was chosen as the primary source for retrieving marketing articles in our current literature analysis. Our objective is to encompass a substantial collection of recently published works on marketing and business intelligence, spanning the years from 2000 to 2023, which totals a 23-year period. As a result, we identified a substantial sample of 1044 articles from various publishers and journals. This diverse selection of publication titles reflects the wide range of topics related to marketing and business intelligence.

Following this initial search, we implemented two additional filtering procedures, labeled as Filters A, B, and C, to enhance the quality of the results and reduce the size of the selected articles. Filter A was applied to limit the selection to articles published in the English language. Filter B was employed to narrow down the results based on the publication year of the articles. Filter C excluded document types such as Notes, Short Surveys, books, and Editorials. The specific search terms and applied filters are as follows:

(TITLE-ABS-KEY("business intelligence") OR TITTLE-ABS-KEY

(marketing)) AND PUBYEAR > 1999 AND PUBYEAR < 2024 AND (LIMIT-TO(LANGUAGE, "English")) AND (EXCLUDE(DOCTYPE, "No") OR EXCLUDE(DOCTYPE, "sh") OR EXCLUDE(DOCTYPE, "bk") OR EXCLUDE(DOCTYPE, "ed"))

In order to identify the most pertinent publications for this research, our primary focus was on pinpointing the most relevant journal articles related to business intelligence within the field of Marketing, covering the past 23 years. We conducted a search encompassing journal article, conference papers, and book chapters (including reviews and articles in press) spanning from 2000 to 2023. This time frame was chosen due to the heightened interest in web-based business intelligence during this period, as noted. To isolate the relevant articles, we executed the following search query: (marketing) AND ("business intelligence" OR "BI"). A brief overview of this collection of 1044 articles is provided in Figures 3 to 6

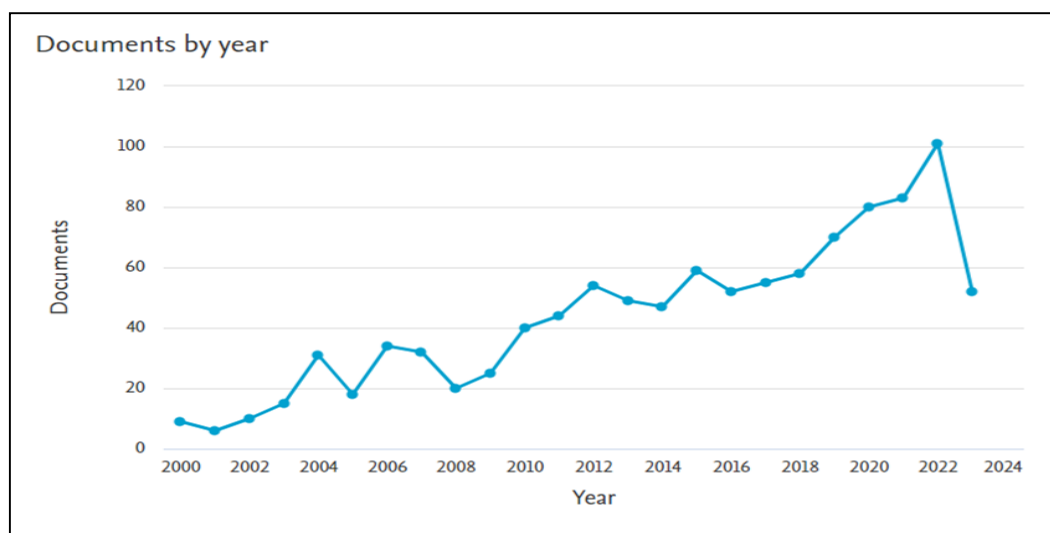


Figure 3. Distribution of articles from 2000 to 2023

An analysis of the publication trends reveals that the number of articles published in this field has been consistently increasing since 2016. Upon initial examination of the

distribution of these articles among various subjects, it becomes evident that despite the query's emphasis on business intelligence in marketing, these papers cover a broad

spectrum of topics. Notably, the majority of papers are focused on Computer Science (n: 510, 25.7%), business and management (n: 325, 16.3%), and engineering (n: 249,

12.5%), as illustrated in Figure 4, which displays the distribution of articles by subject area in the 1044 papers collected.

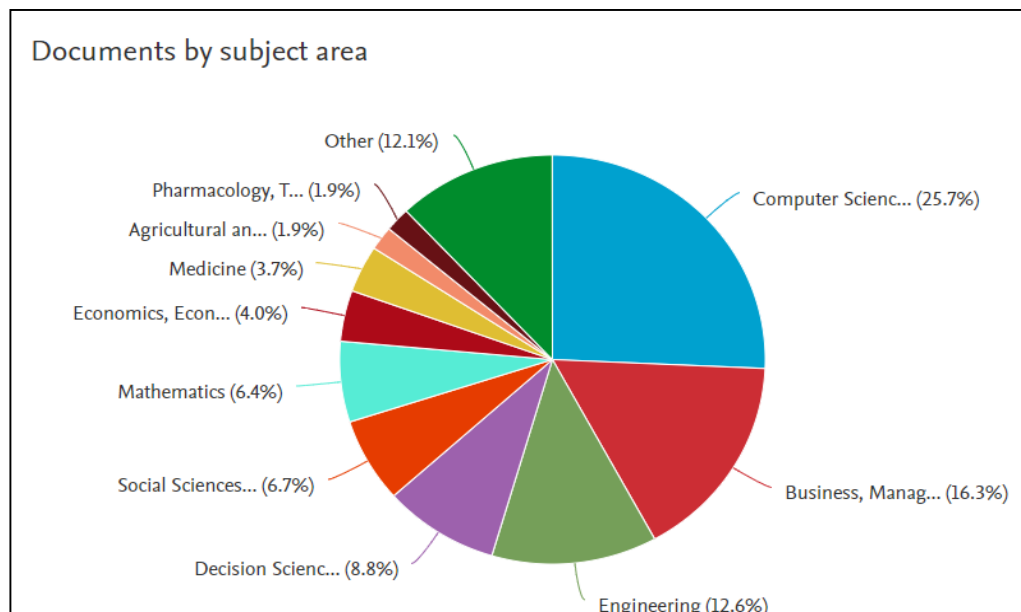


Figure 4. *Distribution of articles from document by subject area*

The breakdown of papers based on their document type reveals that the majority of papers were presented as articles (544 papers, accounting for 51.7%) and conference papers

(310 papers, making up 29.7%). Figure 5 illustrates the distribution of these articles by document type among the 1044 papers collected.

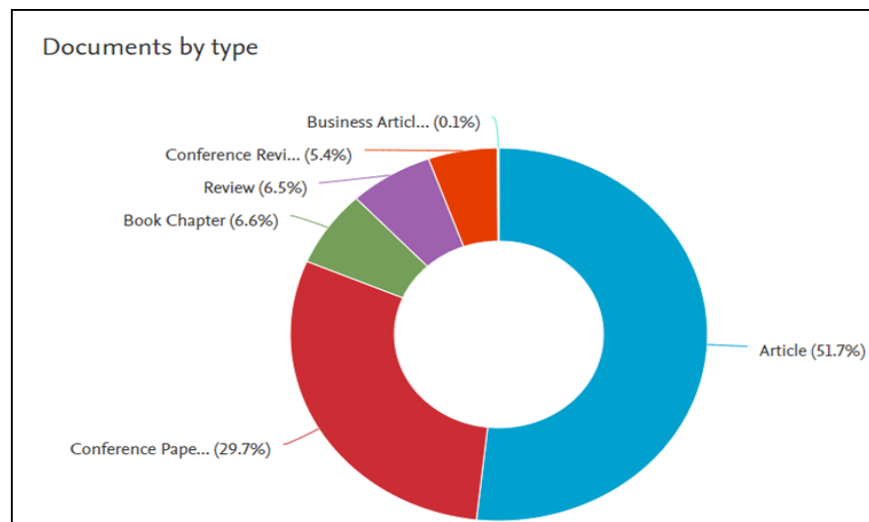


Figure 5. *Distribution of articles from document by type*

In terms of geographical distribution, the papers were predominantly published in the United States, China, India, and the United

Kingdom. Figure 6 provides a visual representation of the distribution of articles by country within the dataset of 1044 papers.

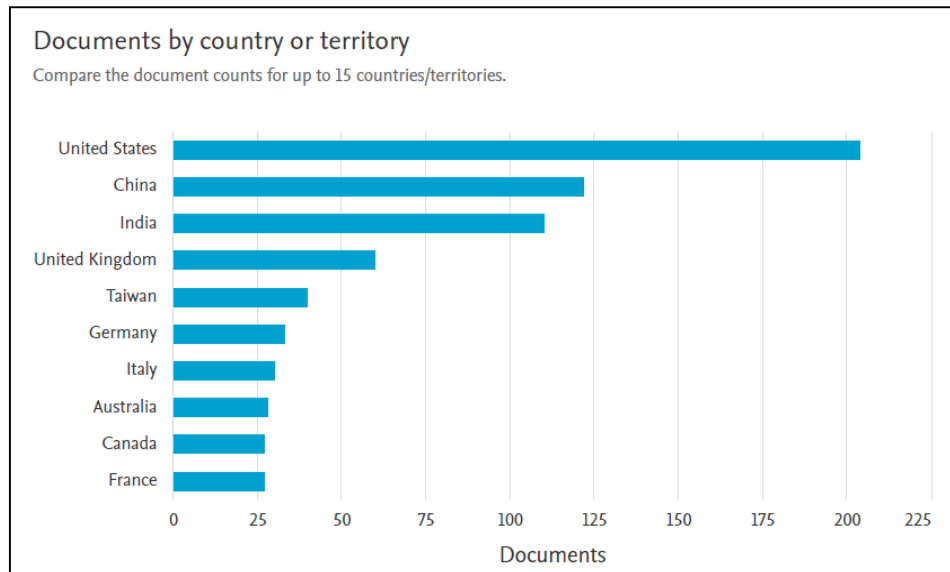


Figure 6. *Distribution of articles from document by country or territory*

Before conducting the analysis, we adhered to established data preprocessing procedures, as outlined in previous studies. These procedures involved the use of Python 3.0 (Python Software Foundation) for data cleaning and the application of word part-of-speech tagging and text processing through the Python library spaCy. During the data cleaning process, we standardized the words in the reviews to lowercase, removed stop words, punctuation, numbers, and non-alphabet characters, and applied stemming to the remaining text (Barrera & Shah, 2023). To enhance the interpretability of the resulting topics and ensure their quality, we restricted the parts of speech of words to "noun" (NOUN), "verb" (VERB), "adjective" (ADJ), or "proper noun" (PROPN). These standard preprocessing procedures play a crucial role in improving algorithm performance and stabilizing the stochastic inference of latent Dirichlet allocation (LDA).

Topic Modeling With LDA

We employed the Latent Dirichlet Allocation (LDA) text mining technique, which suggests that texts are derived from a mixture of different topics (Blei et al., 2003). LDA is known for its efficiency and its ability to produce high-quality topics (Bogoradnikova et al., 2021). From the

dataset we created, we generated two probability distribution results: the distribution of topic probabilities across documents and the distribution of term probabilities within topics (Chen et al., 2021). The number of topics was determined through repeated analyses with different topic counts, comparing the perplexity for each analysis. A lower perplexity value is indicative of a better model fit (Blei et al., 2003), and the perplexity decreases as the number of topics increases. In selecting the optimal number of topics, one must consider both the simplicity and interpretability of the text content.

To evaluate the quality of a particular topic model and determine the optimal number of topics to extract from user reviews, we utilized the coherence score as an evaluation metric. Topic coherence is a qualitative measure used to assess the coherence of a specific topic. It quantifies the semantic similarity between words with high scores within a topic, thereby improving our understanding of the topic's semantic consistency. We made use of the Python package "coherence model" from Gensim to calculate the coherence value. As depicted in Figure 7, the coherence score peaked at a value of 0.45 when the number of topics was set to 5, before gradually decreasing. This indicates that the optimal number of topics

was 10. Following this, we visualized the relationship between these 5 topics and their

associated terms using Python version 3.6.1 and the pyLDAvis tool (Blei et al., 2003).

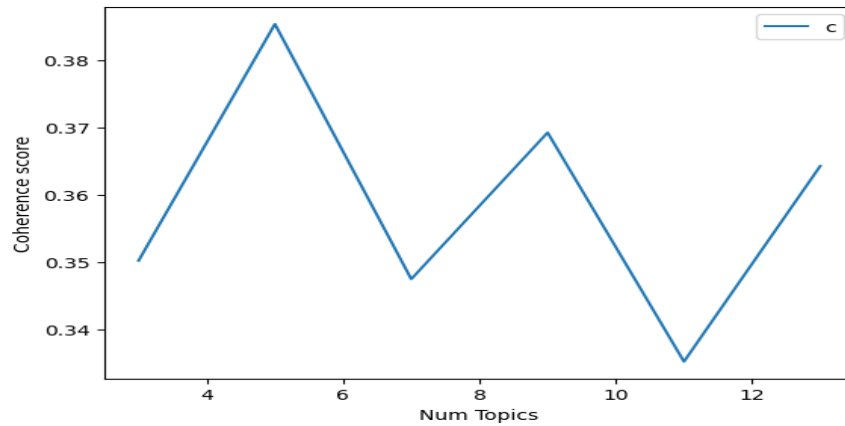


Figure 7. Coherence score for the topic numbers

When λ is set to 1, terms are organized in accordance with their frequency within a topic. As a result, we selected $\lambda=1$ to present the connection between the 5 topics and the top 30 most pertinent terms for each topic, as illustrated in Figure 8. These 5 topics were grouped into distinct themes to enable more insightful analysis, utilizing the computed topic distance (Blei et al., 2003).

circles reflects the topic distance (Blei et al., 2003; Haseli & Sheikh, 2022). The content for each topic was generated using its respective set of keywords (terms) (Haseli & Sheikh, 2022; Haseli et al., 2023). Since the outcomes of statistical measures may not always be straightforward to interpret due to linguistic complexity, we complemented automated text statistics with manual interpretation during the analysis of the topics. The names of the topics were derived from the associated keywords to provide a clear illustration of these topics.

In the 2D representation displayed in Figure 8, the 5 topics are represented as circular shapes. The size of each circle indicates the overall prevalence of the topic, and the distance between the centers of the

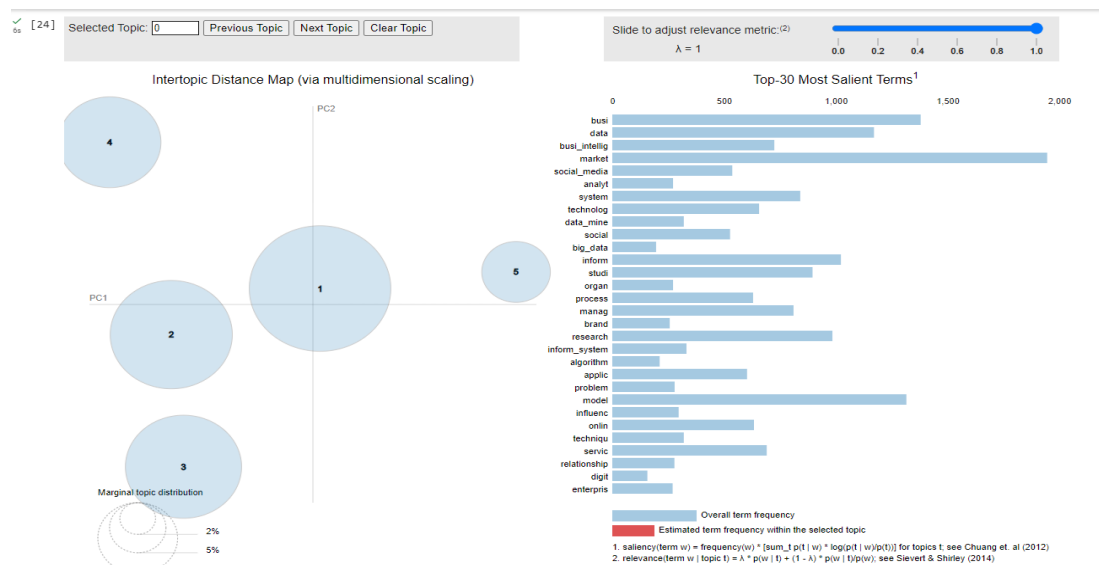


Figure 8. Intertopic distance map and top 30 most relevant terms for topic 5

Results

This research employs the methodology outlined in section 2 to derive a set of topic distributions as mentioned earlier. These outcomes are conveyed in numerical form, making it challenging to readily grasp the features of the topic distribution. In this section, the study utilizes appropriate visual representations to illustrate and scrutinize the findings, with the aim of uncovering specific characteristics within the realm of marketing Business Intelligence research.

Discovering the topics for BI in marketing

In this research, we employ an LDA model to establish five distinct topics, along with the distribution of documents across these topics and the distribution of words within each

topic. We visually represent the connections between topics and words through word clouds, as depicted in Figure 9. Each word cloud corresponds to a specific topic, with the essence of that topic embedded within the words contained in the cloud. Figure 9 showcases multiple word clouds, each comprising the 30 most significant words associated with the respective topic. The size of each word within the word cloud reflects the likelihood of that word appearing within the topic. These words bear a strong connection to research within the realm of Business Intelligence in marketing. Consequently, we can define these five topics based on the semantic content of the words as subfields of research concerning BI in marketing.



Figure 9. Word clouds for Topic 1 – Topic 5.

Topic Prevalence

Using Latent Dirichlet Allocation (LDA) text mining, we categorized the collected data

from 1,044 selected articles focusing on the application of Business Intelligence (BI) in Marketing into 5 distinct topics. (Table 2).

Table 2. Topic classification and keywords

Topic classification	keywords	articles (N=1044), n (%)
Topic 1: business intelligence system marketing planning	Information, business, System, data. Customer, model, management, business intelligence, technology, marketing, base, development, research, analysis, application, service, process, design, product, information system, companies, network, effect, enterprise, new, method, knowledge, framework, online, intelligence	323 (31.5%)
Topic 2: marketing pricing models business	Marketing, product, development, consumer, studies, provide, price, increase, different, data, result, potential, business intelligence, model, change, consider, cost, costumer, treatment, include, need, analysis, offer, available, active, show, one-off, retail, test, level	198 (23.4%)
Topic 3: successful social media marketing brand value	Studies, social media, research, market, effect, online, influence, brand, relationship, impact, service, adopt, consumer, development, role, innovation, communication, model, management, trust, investing, value, social	178 (21.2%)

Topic classification	keywords	articles (N=1044), n (%)
	network, customer, examine, finding, content, perspective, used	
Topic 4: predictive algorithms customer data analytics	Model, problem, algorithm, user, method, propose, base, predict, approach, mine, data, optimize, review, cluster, advertise, interest, set, application, show, product, feature, technique, web, result, recommend, person, extract, task, like, image	146 (16.5%)
Topic 5: big data marketing analytics	Business, data, marketing, analytic, big data, data mining, business intelligence, organization, technology, process, digital, technique, database, retail, help, research, sale, implement, chapter, application, management, tool, store, opportunity, enable, area, business intelligence, need, service, work	55 (7.4%)

Intertopic Distance

Figure 9 presents an overview of the topic model we created. In this illustration, there are five circles that correspond to the five predominant themes found within the 1,044 articles. The distances between these topics are visually represented in a two-dimensional plane using multidimensional scaling, as depicted in Figure 9. It's worth noting that these five topics exhibit significant semantic differences both when compared to each other and when considered collectively.

Topic Terms and Content

Figure 9 also displays the top 30 most pertinent terms for 5 distinct topics. Topic 1, which pertains to "business intelligence system marketing planning," constitutes the largest share, accounting for 31.5% of the tokens, with a total of 323 instances. This topic is chosen as a representative example due to its prevalence. By examining the blue and red bars, which respectively indicate the overall term frequency and the estimated term frequency within the selected topic, we can gain a clearer understanding of the topic's content.

Topic 2, related to "marketing pricing models business," comprises 198 documents, accounting for 23.4% of the total studies. Topic 3, focusing on "successful social media marketing brand value," encompasses 178 documents, making up 21.2% of the studies. Meanwhile, topic 4, dealing with "predictive algorithms customer data analytics," involves 146 documents, constituting 16.5% of the

studies. Finally, the topic labeled "big data marketing analytics" has the lowest representation, with only 55 instances, making up 7.4% of the collected studies.

Topic Classification and Keywords

This section presents findings derived from the analysis of published research on the topic of Business Intelligence in Marketing. The aim of text mining is to identify the most frequently discussed subjects within a collection of literature, rather than conducting an exhaustive review of that literature. Text mining offers three key tools to assist in the process of reviewing a substantial body of research, such as Business Intelligence in Marketing. First, it isolates and ranks the key themes or fundamental elements that form the foundation of this body of work based on their prominence (i.e., how frequently they are discussed). Second, it associates a set of terms or keywords with each theme, also ranked by their prominence. Lastly, it compiles a list of journals and articles that are connected to these themes or their associated keywords. Articles are linked to themes based on the probability of their content being related to a specific theme, given the presence of its most frequently used terms. It's important to note that an article doesn't need to be solely about a theme to be associated with it; rather, the theme's associated terms should be frequently mentioned in the article. Given these capabilities, text mining proves to be a

valuable tool for guiding the reviewing process. Table 1 displays the ranking of themes by their prominence.

Theme 1, "Business Intelligence System in Marketing Planning," examines how the integration of business intelligence systems can enhance marketing efforts. This theme includes thirty correlated terms such as Information, business, System, data, Customer, model, management, business intelligence, technology, marketing, base, development, research, analysis, application, service, process, design, product, information system, companies, network, effect, enterprise, new, method, knowledge, framework, online, and intelligence. The most relevant research papers that elucidate the connections among these terms demonstrate that while many applications of business intelligence systems have proven to be effective in guiding planners to make precise and intelligent business decisions quickly, it is crucial to emphasize that when designing a dashboard for use in an organization with unique characteristics, it should align with their business strategy to maximize the benefits of the business intelligence dashboard. These findings underscore the advantages of business intelligence systems in supporting decision-making and streamlining work processes.

Theme 2, "Marketing Pricing Models in Business," reveals a close relationship between marketing pricing strategies and business intelligence. Pricing models encompass the methods employed to determine the optimal price for a product or service. They play a vital role in how businesses package and present their offerings to consumers. Different pricing models, such as subscription-based or one-off payment models, are available for businesses to employ. Pricing strategies, on the other hand, are internal to a business and assist in identifying the most suitable pricing model to use. Data analytics can enhance pricing models by offering a comprehensive view of a company's sales history, facilitating strategic pricing decisions. When determining the best pricing model to

employ, factors like the product's life cycle, competition, and customer perceptions can be considered. Common pricing models include Product Pricing Model, Ecommerce Pricing Model, and Value-Based Pricing, with the choice depending on the industry and the specific business.

In summary, pricing models influence how goods and services are presented to customers and are external to a business. Data analytics can enhance these models by providing insights into sales history and aiding in strategic pricing decisions. Various pricing models are available to businesses, and the optimal choice depends on the industry and individual business needs.

Theme 3, "Successful Social Media Marketing and Brand Value," establishes a connection between social media marketing and branding. Social media marketing has the potential to enhance a company's brand value or equity, which measures the brand's worth in comparison to others. Effective social media campaigns deliver value to their audience, contributing to increased brand awareness and humanizing the brand. Here are some ways in which social media can boost brand value:

Increased brand awareness: social media has the potential to boost brand recognition, resulting in increased customer retention and a larger customer base willing to support the business.

Direct sales: social media can serve as a novel sales channel, directly influencing the brand's value.

Improved reputation: social media can aid in crafting a more positive brand image, ultimately elevating the brand's worth.

Engagement: Social media enables companies to connect with their customers, address feedback, and cultivate authenticity.

Metrics: Companies can gauge their social media performance by using metrics aligned with their goals, such as return on investment, website traffic, brand recognition, engagement, and lead generation.

It's important to note that while social media alone cannot single-handedly establish

brand value, it can enhance it through various means.

Topic 4 (predictive algorithms for customer data analytics) delves into studies that elucidate customer data analytics. Predictive algorithms in the realm of customer data analytics harness data, statistical algorithms, and machine learning to scrutinize past customer behavior and forecast future actions. These algorithms assist businesses in comprehending customer behavior and taking measures to enhance customer satisfaction and loyalty. Predictive customer analytics entails the examination of prior customer conduct and the prediction of their future actions, utilizing historical data, statistical algorithms, and machine learning techniques. Various predictive analytics models and algorithms, such as clustering models, forecast models, and neural networks, can be employed for customer data analytics. Predictive analytics aids businesses in retaining their current customers and enhancing the customer experience by triggering personalized in-app interactions, adapting in-app communication, and reaching out to customers at risk. While predictive analytics may not be entirely precise, it can unveil trends and patterns that provide data-driven insights into the most effective strategies.

Topic 5 (big data marketing analytics) explores the utilization of big data and marketing analytics to bolster a firm's marketing capabilities. Big data can be harnessed to augment the utilization of marketing analytics, which, in turn, can enhance various aspects of firm marketing, including planning, implementation, brand management, customer relationship management, and product development. The quality of big data marketing analytics (BDMA) has an impact on user satisfaction, perceived value for money, and intentions to reinvest among marketing managers. The relationship between technology and information quality influences market and financial performance, while the extent of deployment affects BDMA technology and information quality. Big data analytics and

managerial support play a vital role in fostering customer satisfaction and CRM capability.

In sum, leveraging big data and marketing analytics can provide firms with a competitive edge in the market. However, it is imperative to consider the quality of the data and analytics in use, along with potential privacy concerns and ethical considerations associated with their utilization.

Conclusions

This analysis of research literature focuses on the utilization of business intelligence in the field of Marketing, with the aim of identifying emerging trends in these applied areas from various perspectives. Marketing is a highly competitive industry known for its innovation, given its vital role in supporting business operations and the economy. This makes it an attractive subject for researchers. Moreover, Marketing generates a substantial volume of data, offering potential benefits from business intelligence applications that can enhance the visibility and recognition of research accomplishments.

To perform this analysis, we gathered a total of 1,044 articles published between 2000 and 2023 from the Scopus database and subjected them to thorough examination. Due to the large number of articles, text mining was employed to better understand the literature. The results revealed a growing interest in business intelligence in Marketing, with each year surpassing the previous one in terms of publication output. Notably, 2016 marked a significant increase in research publications in the analyzed domains. Text mining and topic modeling techniques allowed for a concise summary of the literature by categorizing articles into logical topics defined by key relevant terms. Analysis of author affiliations indicated that most research originates from the United States, China, India, and the United Kingdom. Authors from Asian countries seemed particularly interested in cross-continental research. However, there were relatively few publications from South America and Africa, possibly reflecting a

lower overall interest or research output in business intelligence or data analytics, not necessarily related to the marketing field. Additionally, Computer Science, business and management, and engineering were subjects attracting diverse research interests.

This recent analysis spanned the last 23 years, encompassing the effects of global business transformations on research in this sector. The study provides insights into emerging research trends and potential areas for future research. Through text mining using Latent Dirichlet Allocation, five primary topics were identified, each characterized by the thirty most relevant keywords. These topics revealed several research trends, although the limitations of clustering algorithms like LDA warranted further validation of hypotheses concerning the connections between terms and trends.

The primary conclusion drawn from this analysis is that business intelligence systems continue to dominate research in the marketing industry. Other notable marketing subjects include pricing models, social media marketing's impact on brand value, predictive algorithms for customer data analytics, and big data marketing analytics, mainly focused on enhancing marketing performance. Significantly, in the realm of marketing, business intelligence primarily focuses on prediction rather than modeling and knowledge discovery, underlining the importance of anticipating future developments to improve decision-making.

However, it's worth noting that some studies utilizing marketing problems to evaluate business intelligence techniques may not adequately account for the real business benefits to marketing, as marketing-related terms are often less prominent in those articles. This could be attributed to a lower interest or research output in business intelligence and data analytics in general, rather than a specific issue with the marketing field.

The findings from this study suggest avenues for future research. Firstly, while there is a substantial body of research on business intelligence and marketing, there is

a gap in research specifically addressing the benefits that marketers can gain from business intelligence solutions. As business intelligence adoption in marketing continues to evolve, there is a need to identify the pros and cons of organizations investing in business intelligence for marketing. It is crucial that the implementation of business intelligence solutions aligns perfectly with the unique challenges posed by the business context, making context-aware solutions a priority. This discovery also reveals a research gap in cross-disciplinary research, emphasizing the need for technology researchers to better align the benefits of business intelligence with the field of Marketing.

It is interesting to note that while several marketing-related terms associated with business intelligence and data analysis were included in the analysis (e.g., customer retention, customer segmentation), they were not prominently featured in the topics uncovered. Additionally, some limitations of this study, such as the dynamic nature of business intelligence research and the specific challenges it entails, suggest opportunities for future research to explore the requirements of designing marketing solutions to address these challenges.

References

- Ahmadi, M. M., & Zare, S. (2021). Business Intelligence Technology in Research Organizations (Case Study of Academic Institutes in Tehran). *Journal of System Management*, 6(4), 69-101
- Barrera, K. G., & Shah, D. (2023). Marketing in the Metaverse: Conceptual understanding, framework, and research agenda. *Journal of Business Research*, 155, 113420. <https://doi.org/10.1016/j.jbusres.2022.113420>
- Bogoradnikova, D., Makhnytina, O., Matveev, A., Zakharova, A., & Akulov, A. (2021, May). Multilingual Sentiment Analysis and Toxicity Detection for Text Messages in Russian. In 2021 29th Conference of Open Innovations Association (FRUCT) (pp. 55-64). IEEE. <https://doi.org/10.23919/FRUCT52173.2021.9435584>

- Bremmer, I. (2014). The new rules of globalization. *Harvard business review*, 92(1), 103-107.
- Buenano-Fernandez, D., Gonzalez, M., Gil, D., & Luján-Mora, S. (2020). Text mining of open-ended questions in self-assessment of university teachers: An LDA topic modeling approach. *Ieee Access*, 8, 35318-35330. [10.1109/ACCESS.2020.2974983](https://doi.org/10.1109/ACCESS.2020.2974983)
- Chauhan, U., & Shah, A. (2021). Topic modeling using latent Dirichlet allocation: A survey. *ACM Computing Surveys (CSUR)*, 54(7), 1-35. <https://doi.org/10.1145/3462478>
- Chen, L., Wang, P., Ma, X., & Wang, X. (2021). Cancer communication and user engagement on Chinese social media: Content analysis and topic modeling study. *Journal of Medical Internet Research*, 23(11), e26310. <https://doi:10.2196/26310>
- Chen, Y., & Lin, Z. (2021). Business intelligence capabilities and firm performance: A study in China. *International Journal of Information Management*, 57, 102232. <https://doi.org/10.1016/j.ijinfomgt.2020.102232>
- Daenekindt, S., & Huisman, J. (2020). Mapping the scattered field of research on higher education. A correlated topic model of 17,000 articles, 1991–2018. *Higher Education*, 80(3), 571-587. <https://doi.org/10.1007/s10734-020-00500-x>
- Ghoushchi, S. J., Osgooei, E., Haseli, G., & Tomaskova, H. (2021). A novel approach to solve fully fuzzy linear programming problems with modified triangular fuzzy numbers. *Mathematics*, 9(22), 2937. <https://doi.org/10.3390/math9222937>
- Hadhoud, R., & Salameh, W. A. (2020). How business intelligence can help you to better understand your customers. *International Journal of Business Intelligence Research (IJBIR)*, 11(1), 50-58. <https://doi.org/10.4018/IJBIR.2020010104>
- Haseli, G., & Sheikh, R. (2022). Base criterion method (BCM). In *Multiple Criteria Decision Making: Techniques, Analysis and Applications* (pp. 17-38). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-16-7414-3_2
- Haseli, G., Ögel, İ. Y., Ecer, F., & Hajiaghahi-Keshteli, M. (2023). Luxury in female technology (FemTech): Selection of smart jewelry for women through BCM-MARCOS group decision-making framework with fuzzy ZE-numbers. *Technological Forecasting and Social Change*, 196, 122870. <https://doi.org/10.1016/j.techfore.2023.122870>
- Haseli, G., Sheikh, R., & Sana, S. S. (2020). Base-criterion on multi-criteria decision-making method and its applications. *International journal of management science and engineering management*, 15(2), 79-88. <https://doi.org/10.1080/17509653.2019.1633964>
- Hosseini Astaraei, F., Shojaie, S., Saeidi, P., & Mostaghimi, M. R. (2019). Providing a Model of the Enterprise Market Capabilities with an Emphasis on Organizational Entrepreneurship. *Journal of System Management*, 5(4), 103-112.
- Huang, Z. X., Savita, K. S., Dan-yi, L., & Omar, A. H. (2022). The impact of business intelligence on the marketing with emphasis on cooperative learning: Case-study on the insurance companies. *Information Processing & Management*, 59(2), 102824. <https://doi.org/10.1016/j.ipm.2021.102824>
- Kafilaleh, Y., Bodaghi Khajeh Noubar, H., Motemani, A., & Peyvasteh, A. (2021). Validation of the pattern of brand marketing efforts on social media with customers in the dermato-cosmetic industry. *Journal of System Management*, 7(3), 311-331.
- Kim, K. Y. (2014). Business intelligence and marketing insights in an era of big data: The q-sorting approach. *KSII Transactions on Internet & Information Systems*, 8(2). <https://doi.org/10.3837/tiis.2014.02.014>
- Malek Shirabadi, F., Karimi Zand, M., & Kabaran Zad Ghadim, M. R. (2022). Designing and validating a new integrated digital marketing model. *Journal of System Management*, 8(1), 127-144. <https://doi.org/10.30495/JSM.2022.1944114.1565>
- Kongthanasuwan, T., Sriwiboon, N., Horbanluekit, B., Laesanklang, W., & Krityakierne, T. (2023). Market Analysis with Business Intelligence System for Marketing Planning. *Information*, 14(2), 116. <https://doi.org/10.3390/info14020116>
- Manian, A., & Ronaghi, M. H. (2015). A comprehensive framework for e-marketing implementation by meta-synthesis method. *Journal of Business management*, 7(4), 901-920. <https://doi.org/10.22059/JIBM.2015.57097>
- Moro, S., Cortez, P., & Rita, P. (2015). Business intelligence in banking: A literature analysis from 2002 to 2013 using text mining and latent

- Dirichlet allocation. *Expert Systems with Applications*, 42(3), 1314-1324.
<https://doi.org/10.1016/j.eswa.2014.09.024>
- Moro, S., Pires, G., Rita, P., & Cortez, P. (2019). A text mining and topic modelling perspective of ethnic marketing research. *Journal of Business Research*, 103, 275-285.
<https://doi.org/10.1016/j.jbusres.2019.01.053>
- Peres, R., Schreier, M., Schweidel, D. A., & Sorescu, A. (2023). Blockchain meets marketing: Opportunities, threats, and avenues for future research. *International Journal of Research in Marketing*, 40(1), 1-11.
<https://doi.org/10.1016/j.ijresmar.2022.08.001>
- Rahchamani, S. M., Heydariyeh, S. A., & Zargar, S. M. (2022). A Model for Identification of Factors Affecting Services Intelligent Supply Chains: A Meta-Synthesis Approach. *Journal of System Management*, 8(3), 13-33.
<https://doi.org/10.30495/JSM.2022.1957013.1641>
- Shao, C., Yang, Y., Juneja, S., & GSeetharam, T. (2022). IoT data visualization for business intelligence in corporate finance. *Information Processing & Management*, 59(1), 102736.
<https://doi.org/10.1016/j.ipm.2021.102736>
- Skålén, P., Cova, B., Gummerus, J., & Sihvonen, A. (2023). Marketing-as-practice: A framework and research agenda for value-creating marketing activity. *Marketing Theory*, 23(2), 185-206.
<https://doi.org/10.1177/14705931221123949>
- Sprong, N., Driessen, P. H., Hillebrand, B., & Molner, S. (2021). Market innovation: A literature review and new research directions. *Journal of Business Research*, 123, 450-462.
<https://doi.org/10.1016/j.jbusres.2020.09.057>
- Stone, M. D., & Woodcock, N. D. (2014). Interactive, direct and digital marketing: A future that depends on better use of business intelligence. *Journal of research in interactive marketing*, 8(1), 4-17.
<https://doi.org/10.1108/JRIM-07-2013-0046>
- Tilak, G. (2020). A Review of Using Business Intelligence (BI) in Digital Marketing.
- Vayansky, I., & Kumar, S. A. (2020). A review of topic modeling methods. *Information Systems*, 94, 101582.
<https://doi.org/10.1016/j.is.2020.101582>
- Webster, F. E., & Lusch, R. F. (2013). Elevating marketing: marketing is dead! Long live marketing!. *Journal of the Academy of Marketing Science*, 41, 389-399.
<https://doi.org/10.1007/s11747-013-0331-z>