The Journal of Technology Transfer
https://doi.org/10.1007/s10961-023-10023-z

Leveraging SMEs technologies adoption in the Covid-19 pandemic: a case study on Twitter-based user-generated content

Jose Ramon Saura[1] · Daniel Palacios-Marqués[2] · Domingo Ribeiro-Soriano[3]

Accepted: 13 July 2023

Abstract
The COVID-19 pandemic has caused many entrepreneurs and small and medium enterprises (SMEs) to adapt their business models and business strategies to the consequences caused by the pandemic. In order to identify the main innovations and technologies adopted by SMEs in the pandemic, in the present study, we used a database of 56,941 tweets related to the coronavirus to identify those that contained the hashtag #SMEs. The final sample was analyzed using several data-mining techniques such as sentiment analysis, topic modeling and textual analysis. The theoretical perspectives adopted in the present study were Computer-Aided Text Analysis, User-Generated Content and Natural Language Processing. The results of our analysis helped us to identify 15 topics (7 positive: Free support against Covid-19, Webinars tools, Time Optimizer and efficiency, Business solutions tools, Advisors tools, Software for process support and Back-up tools; 4 negative: Government support, Payment systems, Cybersecurity problems and Customers solutions in Cloud, and and 4 neutral: Social media and e-commerce, Specialized startups software, CRMs and Finance and Big data analysis tools). The results of the present study suggest that SMEs have used a variety of digital tools and strategies to adapt to the changing market conditions brought on by the pandemic, and have been proactive in adopting new technologies to continue to operate and reach customers in a connected era. Future research should be directed towards understanding the long-term effects of these technologies and strategies on entrepreneurial growth and value creation, as well as the sustainability of SMEs in the new era based on data-driven decisions.

Keywords Small-medium sized enterprises · Entrepreneurs · SMEs technologies · Covid-19 · UGC

JEL Classification L2 · L26 · M3 · M15

Extended author information available on the last page of the article

Published online: 29 July 2023

1 Introduction

The COVID-19 pandemic has caused many entrepreneurs and small and medium enterprises (SMEs) to adapt their business models and business structures to the consequences caused by the pandemic (). Both locally and globally, the pandemic has caused economic instability () and changed user purchasing behavior (Metallo et al., 2021; Nurunnabi, 2020). Specifically, due to the situation caused by Covid-19 pandemic, customers have had to modify their consumption habits (Syriopoulos, 2020) in response to the global restrictions proposed by governments and organizations for the control of coronavirus infections (Pedauga et al., 2021). Accordingly, entrepreneurs in their SMEs, have had to make great efforts to acquire and adopt new work methodologies (Juergensen et al., 2020) for their business models and marketing strategies (Guo et al., 2022), digital marketing performance (Haque et al., 2020), communications (Kristinae et al., 2020; Wang et al., 2020), and development strategies.

In this context, as argued by many authors (e.g., Syaifullh et al., 2021;), technological development and innovation (Saura et al., 2023a) have enabled many SMEs to come up with new proposals for sales (Sulistyo, 2016), as well as marketing and business communications. In the context of global connectedness via the Internet and smartphones that have become sources of data and consumption in social networks, SMEs have found novel ways to adapt their businesses to new business forms and formats (Audretsch et al., 2021a, 2021b).

In today's globalized world where technology serves as the common channel of communication through social networks and digital platforms, the emergence of e-commerce for SMEs has also been of a particular relevance (Atanassova & Clark, 2015). At the same time, the market has experienced countless innovations embodied by various digital tools linked to Artificial Intelligence (AI), Machine Learning or Big Data (Pérez-González et al., 2017), as well as other data-centric applications (Belitski & Rejeb, 2022). All these tools allow SMEs to extract insights about their customers and, based on these, adapt their products and services accordingly (Scuotto et al., 2017). Today, global e-commerce is characterized by the 24/7 timeframe in which logistics and competition are accompanied by strategies focused on innovation (Hutchinson & Quintas, 2008) and application of new technologies (Zeng et al., 2010) such as digital marketing (Saura, 2021). Accordingly, companies that get engaged in e-commerce can gain competitive advantages and offer added value to their customers (Chang et al., 2011; Radko et al., 2022).

In this context, and the specific situation caused by COVID-19 (Ribeiro-Navarrete et al., 2021), the present study aims to fill a gap in the scientific literature by using User-Generated Content (UGC) (Daugherty et al., 2008) as a primary source of information to obtain insights regarding technological innovations applied by SMEs during the COVID-19 pandemic (Juergensen et al., 2020; Vivona et al., 2022). UGC is defined as any type of content publicly generated by users on social networks and digital platforms (Borghi & Mariani, 2021). In recent years, UGC has attracted a considerable scholarly attention due to its multiple applications and the possibility it offers in terms of extracting meaningful insights and to create theoretical knowledge related to the society (Chopra et al., 2022; Martínez et al., 2022), economics (Keen, 2011), marketing (O'Hern and Kahle, 2013), finance (Bigne et al., 2021; Tirunillai & Tellis, 2012), tourism (Ray & Bala, 2021), among many other areas.

Likewise, due to the situation caused by the Covid-19 pandemic, it is relevant to understand the main adaptations and the level of resilience shown by the SMEs in the face of the challenge of adapting their business models to a changing ecosystem (Saura et al., 2022a, 2022b, 2022c). In addition, in this context, technologies have increased both the possibilities and the opportunities for success and it is for this reason that we addressed the following main research question (*RQ): What are the primary innovations and technologies adopted by SMEs during the COVID-19 pandemic to adapt their business models based on UGC analysis on Twitter?*

To answer this question, in a database of 21 million tweets related to the coronavirus, we identified those that contained information related to entrepreneurs and SMEs. Next, the data were analyzed using several data-mining techniques (Borghi & Mariani, 2022) including sentiment analysis, topic modelling, and textual analysis. To date, none of the previous studies has applied the aforementioned methodological approach to SMEs and technologies used through innovation in times of pandemic, which justifies the originality of the present study (Runco, 1993) and covers a gap in the literature in terms of applied method development to SMEs adoption of technologies using UGC in the COVID-19 pandemic. The theoretical framework of the present study includes Computer-Aided Text Analysis (CATA) (Short et al., 2010), Natural Language Processing (NLP) and UGC (Marine-Roig et al., 2015). In addition, this study poses the following research objectives: (i) To understand how SMEs have adapted their business models; (ii) To create knowledge about the technologies used by SMEs (Ibáñez et al., 2021) to adapt their strategies to the Covid-19 pandemic; (iii) To explore the opportunities carried out by SMEs and describe the technologies used and (iv) To identify the main business models of SMEs and present their characteristics adapted to the Covid-19 pandemic.

To solve the research problems presented above, we used a database of 56,941 tweets related to the Covid-19 pandemic to identify those that contained the hashtag #SMEs. The final sample was analyzed using several data-mining techniques such as sentiment analysis, topic modeling, and textual analysis. The theoretical perspectives adopted in the present study were Computer-Aided Text Analysis (CATA), User-Generated Content (UGC), and Natural Language Processing (NPL).

The remainder of this paper as follows. Upon a review of the literature in Sect. 2, we discuss the methodology used in the present study in Sect. 3. This is followed by the presentation of the results in Sect. 4. The findings are discussed in Sect. 5. Finally, Sect. 6 draws conclusions and discusses the theoretical and practical implications and limitations of the present study.

2 Literature review

2.1 SMEs and covid-19

Innovation and technologies used by SMEs to overcome the challenges caused by the COVID-19 pandemic have been diverse. SMEs have found in technology the ideal tool to overcome the problems caused by encirclement and mobility restrictions. In addition, innovation has allowed SMEs to modify their business models to make them more flexible and resilient (Menter, 2022).

As demonstrated by Juergensen et al. (2020), European SMEs modified their structures and business models based on the impact and policy decisions taken as a result of COVID-19. Based on different scenarios examined, and depending on the regulation applicable in each case, SMEs were found to focus more on the use of technology for e-retailing or on the innovation of their business models (Lacárcel, 2022).

Similarly, in a study on the impact of the COVID-19 pandemic on the performance of SMEs, Roper and Turner (2020) underscored the importance and relevance of the funding obtained by SMEs from different governmental institutions. The authors also emphasized enthusiasm and ability of SMEs to use new technologies to develop innovation-focused strategies that helped these enterprises to obtain extra resources has previously indicated by Chaithanapat et al. (2022).

At the same time, in a study on innovation in SMEs during the current pandemic, Caballero-Morales (2021) identified the main technology and innovation tools used by SMEs in what is known as “the new normal”. The authors explicitly related risk reduction and economic recovery with SMEs’ willingness and ability to use innovative technological tools in business models (Mariani et al., 2022). Furthermore, in their study on the adoption and state-of-the-art of emerging technologies in SMEs Akpan et al. (2020) investigated the consequences of the emergence of the COVID-19 pandemic focusing on the main methodologies and business strategies used by SMEs to save their businesses (Audretsch et al., 2022a, 2022b). In addition, the authors also identified that cloud computing, Big Data, predictive analytics, and decision-centric innovation have improved the marketing performance of SMEs during these times (Ortigueira-Sánchez et al., 2022; Chopra et al., 2022; Lacárcel and Huete, 2023).

Following this line of argument, Fitriasari (2020) explored the development of commercial techniques used by SMEs to survive during the COVID-19 pandemic. Among other factors, the author found that the most effective tools that SMEs have used to better manage their decisions included the creation of social networks, use of CRMS, opening of new communication channels (González-Padilla et al., 2023), use of Virtual Reality to make remote operations, as well as the Internet-of-Things (IoT).

Similarly, in a study on the technologies and innovation in SMEs, Kamal (2020) highlighted different acquisitions in the use of business models during pandemic under a perspective of the destructive nature of the economy caused by COVID-19. Several authors also highlighted the need for the ethical and sustainable design of SMEs’ operations in times of pandemics. For instance, Kumar et al. (2020) argued that sustainability in operations and the analysis of new challenges for SMEs is a problem that can be solved by using both innovation-centered business models and new technologies (Martín and Fenández, 2022a).

Finally, several previous studies measured the adoption of technologies by SMEs. For example, using the Unified Theory of Acceptance and Use of Technology (UTAUT) model, Kumar and Ayedee (2021) demonstrated the utility and effectiveness of SMEs’ use of technology 4.0, social media, and e-commerce in meeting consumer demands, implementing process automation, and maintaining the imposed social distance restrictions. The authors also emphasized the importance of artificial intelligence for SMEs in their new innovation-focused strategies (Slatten et al., 2021).

These studies demonstrate the ability of SMEs to adapt to change and adopt new technologies. Furthermore, their analysis reflects the interest of researchers in their contributions. Similar studies that use UGC as a primary data source to address research objectives similar to those proposed in the present study are also offered below.

2.2 User-generated content and social media networks

UGC has been consistently used in research in recent decades (Saura et al., 2021a, 2021b). With the rise of social media, the content that users publicly post on social media has become a source of insights and knowledge for companies (Ante, 2023) and governments (Saura et al., 2022a). This type of content can be used to create knowledge and explore emerging themes by understanding users' comments and interactions with each other and with companies as sources of high-quality information. Furthermore, a variety of approaches have been developed to apply filters and generate information from databases containing social media data such as Twitter or Facebook (Saura et al., 2023b).

In this way, authors such as Liu et al. (2017) highlight the study of UGC and its relationship to brands on the Twitter social network. This allows them to identify value creation or users' main perceptions regarding a brand and its products and services (Morstatter et al., 2013). Additionally, authors such as Smith et al. (2012) also focus on this analysis but linked to social networks such as YouTube. Any social network in which content is publicly developed can be analyzed and refined with the ultimate goal of linking the results to the creation of knowledge, using big data tools or AI analysis (Saura et al., 2021a, 2021b).

Also, authors such as Grover et al. (2019), study more traditional models in which the perceived utility or ease of use of a technology such as blockchain for digital transactions is analyzed from the perspective of UGC on the Twitter social network. This content is directly linked to public opinion and therefore, after its categorization and classification, relevant results can be obtained related to the topic chosen for the research. Likewise, authors such as Kar (2021) studied the satisfaction of mobile phone payments by analyzing UGC and creating models based on user communications on social media. Therefore, it can be seen that there are many approaches to understanding the application of UGC in different industries. Furthermore, not only the acquisition or adoption of technology can be studied, but also how information is disseminated through comments or reviews that may allow companies to have direct information from users that can help them improve their products and services in the long term and optimize their offers based on the content that their buyers post on the internet (Susarla et al., 2012). Based on the studies presented, this study uses Twitter UGC to understand comments and information related to SMEs and their adaptation to the new situation caused by COVID-19 in relation to the adoption of new technologies and digital tools (Saura et al., 2023c). Methodology development is presented below.

3 Methodology

In the present study, we followed Fayyad and Uthurusamy (2002) in using data-mining techniques to obtain insights that generate knowledge. As mentioned above, CATA is a theoretical framework used to theoretically validate the results (Palmquist et al., 1997; Pollach, 2012). CATA proposes that the analysis of a text used as a source can be used to build validation, while computational tools are used to make sense of it (Pollach, 2012). In this way, insights can be obtained that create theoretical knowledge for further use by quantitative research that would empirically demonstrate the validity of the results (Berger et al., 2020; Fayyad et al., 1996). Therefore, exploratory approaches can be meaningfully used to create knowledge for future research (Chung & Gray, 1999). According to Krishnapuram and Keller (1996), the use of algorithms that work with machine learning to improve the identification of insights (Cios et al., 2012; Mariani & Borghi, 2021a) is normally used in scientific research to increase the quality of research.

Accordingly, and following Viaene et al. (2001) and Cui and Curry (2005), in the present study, we applied the approaches focused on machine learning to increase the quality of the obtained insights and to efficiently highlight the results based on automation and data mining protocols (Chae, 2015). For the development of the three steps tha compmose the indicated methodology, we used Python to perform sentiment analysis, topic modeling and textual analysis approaches.

Specifically, we relied on Support Vector Machines (SVM) (Hilmersson et al., 2021) and Support Vector Classifillers (SVC) (Kim & Sohn, 2010) presented and analyzed below (Grover et al., 2018). In addition, a topic modelling algorithm was used to extract topics in the form of quantitatively represented clusters of nodes (Ramage et al., 2009). Finally, textual analysis software and Python were applied to extract insights in relation to the proposed objectives (Isoaho et al., 2021). NLP is a subfield of linguistics that, in combination with computer sciences and AI attempts to extract knowledge from the interactions between humans and people through language (Chowdhury, 2003; Mariani & Borghi, 2021b).

With the use of specific software for the analysis of these interactions, the results might provide a deeper understanding of the connotations, ideas, arguments, or communications that exist between the language and the digital platform where such content is created (Cai, 2021). NLP, therefore, is directly linked to UGC (Ray et al., 2020). Indeed, as argued by Liu (2020), the combination of both techniques can determine the robustness and strength of the structure of both a database and the techniques used to extract insights from it (Krippendorff, 2018). In this context, social networks are deemed to be are relevant data source for research (see also Lai and To 2015; Warner-Søderholm et al., 2018). In summary, the conceptualization, analysis, and processing of data with NPL (Mariani & Borghi, 2020; Nadkarni et al., 2011) and UGC techniques allow researchers to identify indicators, variables, or relevant information that can facilitate future research.

3.1 Data collection, processing and sampling

In the present study, we used a sample of 21 million tweets containing information related to coronavirus (e.g., with hashtags #COVID-19, #Coronavirus) collected between March 20, 2020, and February 15, 2021. The tweets were downloaded daily from the Twitter API for 12 months until further filtering (Tufekci, 2014). Python and Pandas libraries were then used to filter and debug the tweets. In the process of debugging and filtering the databases, duplicates and repeated tweets were removed (Lipizzi et al., 2015). The name of the users, the ID of the tweets, the profile descriptions of the users, the URLs of the Tweets, or the accompanying images were excluded, as only the text of the tweets was part of the study that was analyzed from the perspective on NLP (Wolfe et al., 2021). Of the total 21 million tweets, a total of 62,029 were tweets related to entrepreneurs and SMEs as we used this dataset to serch the following terms "SME", "SMEs" and "Small and medium-sized enterprises". After the removal of duplicates (n = 5,088), the final sample including those terms was reduced to 56,941 tweets (Saura et al., 2022a, 2022b, 2022c).

3.2 Textblob sentiment analysis

In order to create knowledge and identify insights from the obtained sample obtained, we used both approaches focused on user-generated content on Twitter and textual analysis focused on CATA analysis (Meyners et al., 2013). The approach focused on sentiment analysis followed the theorization of content analysis proposed by Nemes and Kiss (2021) and is linked to NPL. Specifically, we used Textblob, a process and research method for the development of sentiment analysis (Laksono et al., 2019). A library developed in Python for text processing was used (Hasan et al., 2018). This algorithm was built on NLTK (Hardeniya et al., 2016). The approach was validated by four experiments focused on the development of sentiment analysis with SVC, multi nominal naive, logistic regression, and random forest classifiers (Hiremath & Patil, 2020). Therefore, despite the limitations of sentiment analysis, we increased the quality of our results by avoiding incidences that can be related to connotations, sarcasm, irony and other language features (Bender & Friedman, 2018; Mariani & Borghi, 2022).

The results were expressed in three numerical ranges (Hiremath & Patil, 2020). First, polarity was expressed as -1, while subjectivity was expressed as 0.0 or 1.0. The algorithm was trained a total of 741 times, with manually classification of the tweets upon reading them. The algorithm working with machine learning was trained to learn on its own, and the more tweets it analyzed, the higher was the percentage of accuracy efficiency it obtained. In the next step, the tweets were divided into three databases (each containing the tweets that expressed negative, positive, and neutral sentiment), and a topic-modeling LDA (Latent Dirichlet allocation) algorithm was used to extract relevant topics from the data (Mariani & Matarazzo, 2021).

The classification of data into subsets expressing three sentiments is a standardized mechanism in this type of sentiment analysis research (Ahmad et al., 2017; Duran et al., 2010). The results were then validated with the indicators of accuracy, recall, f1-score and support (Shokhnekh et al., 2019). Subsequently, the best results of the four experiments were selected to start the next phase of the methodological approach process (Onan et al., 2016).

3.3 Topic modeling with latent Dirichlet allocation

As the next step, we used the LDA algorithm (Maier et al., 2018). This mathematical algorithm is a standard application of the extraction of insights from large databases in the scientific literature. This type of algorithm allows for the analysis of words and their organization in large databases (Yu et al., 2019). In the present study, the documents were proposed as clusters of tweets, in order to identify the themes that compose them according to the sentiment detected in the previous phase (Tajbakhsh & Bagherzadeh, 2019).

Once the algorithm was applied to the databases, indicators regarding the number of words that make up each database were obtained (Zhao et al., 2011). Based on the most frequently used words in the databases, we labelled each of the identified topics (Tajbakhsh & Bagherzadeh, 2019).

In general, the LDA algorithm is a probabilistic assumption developed in different phases. In the first phase, the documents are separated according to their weight

(Xue et al., 2020); the second phase, they are distributed into topics, and the most repeated words are grouped by the algorithm (Yang & Rim, 2014). Once the keywords are obtained based on frequencies and groupings, the obtained topics can be assigned names. With the objective of identifying the most salient topics in linguistic terms, the first 20 topics generated by the LDA algorithm are examined. Subsequently, utilizing a semantic lens, rather than a statistical one, the most appropriate topics are selected for inclusion in the present study based on their relevance to the indicated research question and objectives (Mariani & Baggio, 2022).

3.4 Textual analysis

The final stage of our analysis included the application of textual analysis where the weight of the main keywords in the analyzed databases was taken into consideration (Carley, 1994). At this step of the process, databases of tweets expressing different sentiments and topics were already obtained based on the frequency and weight of the keywords in those databases, which allowed us to specifically indicate and shape meaningful insights (Barbier et al., 2012; Saura et al., 2021a, 2021b).

In principle, textual analysis is directly linked to the content analysis proposed by Krippendorf (2018). For instance, Ceci and Lubatti (2012) highlighted the use of textual analysis and content analysis in SME social networks. These analysis techniques were also used by Anand et al. (2021) in their research on the evolution, background, and future directions of SMEs to share knowledge and transfer it to this industry.

In the present study, following and Kobrinskii et al. (2018), we identified the most frequently repeated words and used unigrams and bigrams (Ghiassi et al., 2013) to understand relevant indicators (Hassan et al., 2020). Textual analysis is based on theories such as Mutual Information (MI) (Datta et al., 2017). MI is a statistical measure that focuses on the concurrence of words and analyzes them to find meaningful correlations (Saura et al., 2021a, 2021b). Using CATA and textual analysis, one can identify the frequency of words grouped in nodes or groups of words and, based on their statistical indicators, to define the main insights linked to the object of the study (Barbier et al., 2012).

4 Results

4.1 Textblob sentiment analysis results

As discussed in Sect. 2, in the present study, we used the Textblob sentiment analysis algorithm (Kaur & Sharma, 2020) that works with SVM (Wang, 2005) and SVC (Lee & Lee, 2007) known as logistic regression, naïve Bayes, SVC, and random forest classifier (Banker & Patel, 2016). The measures that were previously found to be more successful in terms of accuracy were used in the present study. As argued by Haddi et al. (2013), accuracy is one of the most used metrics in the studies developed with SVM and SVC to obtain precision in the results. Therefore, the higher the percentage of accuracy of the study, the greater the capacity for accuracy and possible prediction of the results (Cao et al., 2013). In our data, the highest results in terms of accuracy were obtained for Linear SVC Sl. No. 8 (0.992727) and 18 (0.958909) relative to Logistic Regression results (see Table 1).

Table 1 Model category details of Textblob accuracy

Sl. No	Model Name	Fold_idx	Accuracy—Textblob
0	RandomForestClassifier	0	0.609960
1	RandomForestClassifier	1	0.613232
2	RandomForestClassifier	2	0.614686
3	RandomForestClassifier	3	0.618909
4	RandomForestClassifier	4	**0.622182**
5	LinearSVC	0	0.965104
6	LinearSVC	1	0.977826
7	LinearSVC	2	0.981461
8	LinearSVC	3	**0.992727**
9	LinearSVC	4	0.969455
10	MultinomialNB	0	0.860051
11	MultinomialNB	1	0.862232
12	MultinomialNB	2	**0.900763**
13	MultinomialNB	3	0.886182
14	MultinomialNB	4	0.898909
15	LogisticRegression	0	0.937114
16	LogisticRegression	1	0.942930
17	LogisticRegression	2	0.954925
18	LogisticRegression	3	**0.958909**
19	LogisticRegression	4	0.953455

Table 2 Summarized brief scores by model name and Textblob analysis

Sl. No	Model Name	Scores of Textblob analysis
1	LinearSVC	0.992727
2	LogisticRegression	0.958909
3	MultinominalNB	0.900763
4	RandomForestClassifier	0.622182

In addition, to better understand the overall results of Textblob sentiment analysis, Table 2 presents the values as a set of accuracy in relation to the four experiments performed in the present study. This is a standard process used in applying machine learning and content analysis algorithms (Shalev-Shwartz and Ben-David, 2014). In this way, the results that do not yield the expected accuracy can be discarded, the study process can be continued with the models that yielded a higher accuracy (Ayodele, 2010).

Once the sentiment results were obtained, a classification report was developed in order to understand the total metrics (Kaur & Sharma, 2020). Table 3 shows the results of the classification report for sentiment analysis for the three different sentiments (positive, negative and neutral), as well as the type of accuracy obtained for each of them. Table 3 also lists the recall, fl-score, support values and accuracy variables. According to Shalev-Shwartz and Ben-David (2014), accuracy is a metric that indicates the quality of measurement and processing of the algorithm used in machine learning for the classification

Table 3 Classification report for sentiment analysis with Textblob

Sl. No	Parameters	Vader			
		precision	recall	f1-score	support
1	Negative	0.30	0.00	0.00	79
2	Positive	0.41	0.09	0.15	320
3	Neutral	0.32	0.90	0.47	183
4	Accuracy			0.33	582
5	Macro avg	0.24	0.33	0.21	582
6	Weighted avg	0.33	0.33	0.23	582

process of the assigned tasks. In our case, the task was the division according to sentiment expressed in the analyzed tweets (Banker & Patel, 2016).

Furthermore, the recall indicator reflects the quantity that the machine learning model can identify in a database (Cao et al., 2013). The F1-score indicator combines precision and recalls to be defined as a single value that facilitates the comparison of the global model based on accuracy (Ayodele, 2010). Following Kaur and Sharma (2020), the support indicator measures the amount of support needed from machine learning to predict the outcome of the model in the form of sentiment. Finally, the macro average indicator measures the average total of the model based on the analyzed indicators and the weighted average measures its relativity in terms of weight (Wang, 2005). Table 3 shows the results for each indicator based on the proposed sentiments.

As can be seen in Table 3, the highest precision value was obtained for positive sentiment tweets (0.41), neutral sentiment tweets (0.32), and negative tweets (0.30). Of the total sample, 57.71% tweets were classified as expressing positive sentiment, 12.83% as expressing negative sentiment, and the remaining 29.45% as expressing neutral sentiment. Furthermore, Table 4 shows the polarity and subjectivity variables identified in our experiments.

4.2 Latent Dirichlet allocation analysis results

Through the use of the LDA model (Xue et al., 2020), a total of 15 topics related to the technologies and innovations made by SMEs in the COVID-19 pandemic were identified. For their identification, LDA was applied to the set of databases extracted from the sentiment analysis process. As the primary focus of the present study was on technologies or innovation actions, the irrelevant topics identified in the data were eliminated from the results (Zhao & Jiang, 2011). For the formulation of topic names, the last 10–20 words obtained with the LDA were used (Ostrowski, 2015). Specifically, the LDA was applied independently on each of the databases divided into three sentiments.

In order to statistically quantify the results and the relevance of the analyzed topics, the keyness and p-value values were calculated. These indicators were computed to measure the strength of the link between the topics and to determine the log-likelihood score values to assess the relevance of the topic in the total database (Rayson & Garside, 2000). The log-likelihood value > 3.8 was deemed to be statistically significant. Furthermore, p-value < 0.05 was considered statistically significant. Table 4 presents a

Table 4 Polarity and subjectivity analysis

Parameters	Average subjectivity	Average polarity
Negative	0.365301	-0.151567
Positive	0.508248	0.295351
Neutral	0.006919	-0.002567

summary of the identified topics identified together with their description, sentiment, as well as keyness and *p*-value values (Table 5).

4.3 Textual analysis results

Next, we performed textual analysis to identify the weight of keywords in the overall database, the number of keywords related to the object of study, and the weight of certain groups or word nodes with respect to the total of the study. To this end, we applied NLP techniques the UGC database of tweets using GroupBy in Python (Bhavsar & Manglani, 2019). The meaningful variables related to the words' frequency (Fq.) and weighted percentage (WP) (Krippendorff, 2018) are presented in Table 6. Following Kaur and Sharma (2020), we identified similar words linked to the previously identified topics to understand what types of keywords make up each topic, as well as their weight within each topic. It should be noted that although there are other words in the topic, the variable Fq. indicates those words selected in relation to the textual analysis process for those words selected as keywords to define each topic.

Finally, in order to understand how words are organized and classified in the analyzed databases, a Python n-gram study, a standard process in CATA-based research, was applied. An n-gram model predicts the occurrence of a word on the occurrence of its n—1 previous word (Short et al., 2010). Likewise, a bigram (n = 2) predicts the occurrence of a word only its previous words as n-1 = 1 in form of collocates. Table 7

Table 5 Identified topics and measurements

Topics		Sentiment	Keyness	*p*-value
1	Free support against Covid-19	Positive	520.28	0.025
2	Government support	Negative	478.02	0.021
3	Webinars tools	Positive	466.60	0.018
4	Social media and e-commerce	Neutral	391.97	0.017
5	Time Optimizer and efficiency	Positive	201.24	0.012
6	Business solutions tools	Positive	182.01	0.011
7	Advisors tools	Positive	174.52	0.011
8	Software for process support	Positive	120.39	0.010
9	Customers solutions in Cloud	Negative	99.51	0.009
10	Cybersecurity problems	Negative	91.05	0.008
11	Specialized startups software	Neutral	86.03	0.008
12	CRMs and Finance	Neutral	81.01	0.006
13	Big data analysis tools	Neutral	79.83	0.006
14	Back-up tools	Positive	79.71	0.006
15	Payment systems	Negative	78.49	0.006

Table 6 Textual analysis results by a group of keywords

Tp	Words in topic	Fq	WP
1	SME policy, Supporting SMEs, SMEs pandemic competition, SME free online platforms, etc	2091	16.97
2	SME government reponses, support resources, support Hub, government measures, among others	1204	15.64
3	Webinar for SMEs, COVID-19 webinars and SMEs, SME's perspectives, resilience agains Covid-19	780	13.28
4	Social media adoption, SMEs media adoption, online strategies, measures in response, Covid-19 crisis, Actively used, digital solutions in times of COVID-19, among others	652	12.23
5	Time optimization, management efficiency, real-time optimization, reduced cycle times, monitoring, etc	578	11.17
6	Business solutions, project management, management software, data solutions, management systems, etc	502	10.95
7	Advisors tools, checklist advises, capacity, services program, local advisory, among others	432	9.42
8	Software for process support, entrepreneurship policy responses, tech support, business management, etc	427	9.41
9	Customers solutions in Cloud, fast-tracking cloud, cloud technologies, economy recovery, etc	402	9.39
10	Cybersecurity problems, cyber threats, digital security, practical tips, recommendations, among others	341	9.25
11	Startups software, opportunities for startups, data startups software, startups support solutions, etc	320	9.09
12	CRMs and Finance, CRM resources, economic growth, financial advise, economic flow, among others	286	8.37
13	Big data analysis, data-driven solutions, data analysis, prediction software, recovery analysis, etc	189	7.37
14	Back-up tools, practical help, recovery priorities, prioritise the data, recovery systems, among others	173	7.28
15	Payment systems, digital payments, payments solutions, online payments, among others	162	7.05

Table 7 N-grams for the collocates of the positive topic "Free support against Covid-19"

R	Collocates for IoT			
	Freq	Freq L	Freq R	Collocate
1	1750	791	959	Freesupport
2	842	429	413	Help
3	301	142	159	Measures
4	261	146	115	CalltoAction
5	109	52	57	Responses

Table 8 N-grams for the collocates of the negative topic "Government support"

R	Collocates for IoT			
	Freq	Freq L	Freq R	Collocate
1	1207	589	618	Financinsupport
2	752	351	401	Impactedbusinesses
3	283	159	124	Commonmeasures
4	125	57	68	Negativeeffects
5	114	48	66	Recovery

Table 9 N-grams for the collocates of the neutral topic "Specialized startups software"

R	Collocates for IoT			
	Freq	Freq L	Freq R	Collocate
1	301	156	145	Software
2	294	129	165	Startupchallenges
3	153	78	75	Programs
4	124	59	65	Solutions
5	89	42	47	Startupopportunities

shows the division the different n-grams based on the supported placement, both to the left and to the right of the words under study.

This approach makes it possible to understand the context of the word or similar words in the database and make decisions regarding its meaning in the process of corpus understanding and linguistic computation (Biber, 2004; McEnery & Hardie, 2013). Tables 7, 8, 9 present the topics with the highest p-value by sentiment in the form of n-grams linked to their weight in the database according to Rank (R). The insights from the extracted information are discussed in Sect. 5.

5 Discussion

As mentioned in Sect. 1, in the past year, SMEs have suffered from the consequences of the COVID-19 pandemic and had to adapt their business and strategies to a new ecosystem that prioritizes economy and mobility restrictions (Caballero-Morales, 2021). As argued by Fitriasari (2020), this fact has prompted many SMEs to open new ways of developing and using tools that could help their businesses to counteract this situation (Guo et al., 2022). Accordingly, innovations, communication strategies, and solutions that focus on the utilization of novel technological processes have elicited growing interest within the scientific community (Haque et al., 2020), particularly as they pertain to SMEs. In this context, the present study identified topics linked to sentiments that, in turn, highlight the use of both tools and new innovation-focused strategies performed by SMEs in their businesses.

Furthermore, the positive Topic 1, "Free support against COVID-19", has obtained the keyness of 520.28 and *p-value* of 0.025. Various initiatives in relation to support for small and medium-sized enterprises have been highlighted as sources of information for further research. This includes programs such as emergency loans, tax relief and grants to help businesses cope with the economic impacts of the pandemic (Haque et al., 2020). The motivations of the support, as well as the actions taken to boost the SME sector in the pandemic, reflect the solidarity and social support of companies from different industries. These efforts are not only aimed at helping businesses survive, but also at maintaining jobs and supporting the overall economy. This is consistent with the findings of which highlights the important role of SMEs in the economic recovery efforts. Additionally, it is important to note that further research is needed to understand the long-term effects of the support provided to SMEs and the impact it has on the overall economic recovery (Bui and Lo 2022).

The use of digital platforms and tools focused on government support was highlighted in the negative Topic 2, "Government support" (keyness = 478.02; *p-value* = 0.021). Similarly, several previous studies also highlighted negative sentiments in this type of processes and strategies to support SMEs, particularly in light of the challenges posed by the coronavirus pandemic such as late payment, insufficient support, and inadequate financial aid. These findings align with studies by Nemes and Kiss (2021) and Nurunnabi (2020) which noted similar issues facing SMEs during the pandemic. It is important to acknowledge and address these negative sentiments in order to improve the effectiveness of support measures for SMEs. Furthermore, it is crucial for governments and policy-makers to take a comprehensive and holistic approach in addressing the challenges faced by SMEs, in order to ensure their survival and economic recovery during and after the pandemic (Martín et al., 2022).

Topic 3 that was identified in the data was "Webinar tools" (keyness = 466.60; *p-value* = 0.018). This topic was found to be associated with positive sentiment. Similarly, Pedauga et al. (2021) investigated how, in the pandemic period, SMEs adopted technologies in relation to communication and online meetings (Syaifullah et al., 2021), which has triggered the development of new business and marketing strategies by SMEs (Syaifullah et al., 2021). Furthermore, the increased use of webinar tools has allowed SMEs to adapt to the changing market conditions and continue conducting business remotely. This shows that webinar tools have become essential for SMEs in the context of the COVID-19 pandemic, as they enable remote communication and support for virtual meetings, which is important for the continuity of business operations.

Furthermore, Topic 4, "Social media and e-commerce" (keyness = 391.97; *p*-value = 0.017) highlights the importance of digital transformation for SMEs in the current environment. The use of social media platforms and e-commerce solutions have become essential for businesses to continue to reach and engage with customers, as well as to sell goods and services in a digital environment. Additionally, this topic also highlights the challenges faced by SMEs in terms of adapting their business models and strategies to the digital landscape. According to research conducted by Scuotto et al. (2017), SMEs have had to implement new marketing and communication strategies that involve the use of digital tools, such as social media platforms, to maintain their competitiveness and reach out to new customers (Khatami et al., 2022).

Furthermore, Topic 5, "Time optimizer and efficiency" (keyness = 201.24; *p-value* = 0.012), highlights the use of tools, usually with a characteristic linked to process innovation, to efficiently measure and optimize the actions performed by SMEs (cf. Kamal, 2020). In addition as indicated by Topic 6, "Business solution tools" (keyness = X182.01; *p-value* = 0.011). SMEs have realized the importance of adopting and utilizing tools focused on business solutions, which are typically geared towards the customization of services and products that are closely tied to the actions performed by the SMEs. These tools can help SMEs stay competitive and meet the changing needs of their customers (Kristinae et al., 2020). Furthermore, the utilization of these tools can also increase efficiency, productivity and profitability of SMEs in their operations (Laguía & Moriano, 2021). For example, this may include software applications that allow for better communication, customer relationship management, and financial management. Through the integration of these tools into their processes, SMEs can streamline their operations and make better-informed decisions. Therefore, the adoption of these tools for SMEs can be essential to their survival in today's digital age (Roper & Turner, 2020; Scuotto et al., 2017).

With respect to Topic 7, "Advisors tools" (keyness = 174.52; *p-value* = 0.011), this topic was found to be associated with a positive sentiment. According to this topic, SMEs have made use of tools and applications as well as platforms to obtain advice from experts who, as a sign of solidarity, offered their services free of charge to compensate for the losses caused by the pandemic (Nurunnabi, 2020). In relation to Topic 8, "Software for process support" (keyness = 120.39; *p-value* = 0.010), SMEs used the tools focused on process support, Similarly, Juergensen et al. (2020) already indicated that, in situations of uncertainty, companies can adapt their businesses by using tools that meet the new needs (Zeng et al., 2010). The use of these tools and platforms by SMEs can be seen as a means of adaptation to the changing market conditions caused by the pandemic, and can also be viewed as a strategic decision to ensure the continuity of business operations. Furthermore, the use of these tools and platforms also reflects the importance of digitalization in SMEs, and the need to adopt new technologies and digital solutions in order to stay competitive in today's market. Additionally, the use of software for process support can help SMEs to streamline their operations and improve efficiency, ultimately leading to cost savings and improved performance.

However, Topic 9, "Customers solutions in the cloud" (keyness = 99.51; *p-value* = 0.009) was found to be associated with negative sentiment. This topic suggests that the adoption of technologies by SMEs is not a simple matter, since there are different business models and strategies that must be appropriately adapted depending on the knowledge of the entrepreneurs (Wang et al., 2020). Innovation through the cloud has impeded SMEs, as innovation requires both knowledge related to its installation and adaptability to the way companies work, as well as daily management of such technology (O'Kane et al., 2021; Xue et al., 2020). As argued by Wang et al. (2020), it is important to note that the adoption of technology by SMEs is not a one-size-fits-all approach. Different businesses have different needs and capabilities, and it is crucial for entrepreneurs to have a thorough understanding of these factors in order to effectively adopt and implement new technologies. As discussed before, the adoption of cloud-based technology is not a straightforward process and requires not only technical knowledge but also an understanding of how it can fit into the specific business operations and daily management. As highlighted by Saura et al. (2022a), SMEs need to have the technical and management expertise to fully leverage the benefits of cloud-based technologies.

In addition, as a consequence of the increased use of innovation and digital strategy (Heredia et al., 2022), the negative Topic 10, “Cybersecurity problems” (keyness = 91.05; *p-value* = 0.008), highlighted the difficulty for SMEs related to protecting the data created as a result of digital strategies. Thus, cybercriminals have developed new strategies that target vulnerabilities related to the installation and suitability of connected devices and digital platforms adopted by SMEs. This can lead to problems, data leaks, and extortion (Akpan et al., 2020). Likewise, the adoption of these technologies also requires a proper understanding of the potential risks and challenges that come with them, such as cyber threats and data breaches. Therefore, SMEs must also implement adequate security measures and protocols to protect their digital systems and data. SMEs need to develop a comprehensive cybersecurity plan that includes regular employee training, network monitoring, incident response plans and regular testing and updating of security systems to protect against potential threats. It is also important for SMEs to collaborate with expert service providers to help them identify vulnerabilities and detect potential cyber threats, to reduce their risk of falling victim to cyber-attacks (O’Kane et al., 2021).

Furthermore, as indicated Sulistyo (2016), initiatives developed by startups have been consolidated with specific resources and solutions for SMEs, giving rise to Topic 11, “Specialized software startups” (keyness = 86.03; *p-value* = 0.008), which concerns solutions for SMEs to perform their duties during the pandemic (Anand et al., 2021). Therefore, Topic 11, highlights the importance of specialized software startups as a means of support for SMEs in navigating the challenges posed by the pandemic. In particular, these startups have been seen to provide valuable tools and resources to help SMEs adapt to the new business environment and continue operating effectively (Yi et al., 2022). Additionally, the growth of these startups is expected to continue as SMEs look for new ways to streamline their processes and improve efficiency in order to stay competitive. Sulistyo (2016) suggest that specialization in software solutions can lead to better results by SMEs as they offer targeted services, reducing the risk of ineffective solutions for SMEs. Also, as the use of software for SMEs increases, it is important for governments and other organizations to provide training and education for SMEs to understand these software better, and to maximize the benefits for their businesses (Barbosa et al., 2022).

Also, Topic 12, entitled “CRM and finance” (keyness = 81.01; *p-value* = 0.006), suggests that the adoption of technologies linked to the management and processing of data, as well as the organization of accounting and taxation of SMEs, has been a recurrent option during the COVID-19 pandemic (Pérez-González et al., 2017). Likewise, this topic highlights the importance of CRM systems and financial management tools for SMEs during the pandemic. These systems enable SMEs to efficiently manage their customer interactions and financial data, providing them with valuable insights into their business operations and helping them to make informed decisions (March-Chordà et al., 2021). Additionally, the integration of these systems with existing enterprise resource planning (ERP) and accounting software can enable SMEs to automate and streamline their business processes, resulting in improved efficiency and cost savings (Pérez-González et al., 2017). The use of this software enables SMEs to have a better understanding of their financial position, to make more accurate forecasting, and to take advantage of the possibilities of the digital age to adapt to the market changes and to be more competitive (Saura et al., 2022a).

Also, Topic 13, “Big Data analysis tools” (keyness = 79.83; *p-value* = 0.006), highlights the adoption of technologies focused on data analysis by SMEs (Hutchinson & Quintas, 2008). This point was also highlighted in previous research on the value of data for small businesses (Hutchinson & Quintas, 2008; Kumar et al., 2020). As indicated by Saura et al.

(2022b) SMEs may also benefit from using Big Data visualization tools to better understand their data and make data-driven decisions. These tools can help SMEs identify trends and patterns in their data, allowing them to make informed business decisions and stay competitive in the market (Rivna & Gress, 2022). Furthermore, the use of advanced analytical techniques such as machine learning and AI can also help SMEs gain valuable insights from their data, and make predictions about future market trends and customer behaviors. However, it is important for SMEs to have a clear understanding of the data they have, and the data they need, as well as the capabilities and limitations of the tools they use, in order to effectively analyze their data (Ribeiro-Navarrete et al., 2021).

Likewise, Topic 14, "Backup tools" (keyness = 79.71; *p-value* = 0.006), emphasizes the interest of SMEs in using and adopting tools that help them to have backups. This is interesting because, in times of crisis and uncertainty, backups can help companies to make better decisions and solve problems that may be caused as a result of new processes (Kumar & Ayedee, 2021). As SMEs face an increasingly digitalized and interconnected business environment, the importance of data security and protection becomes crucial. Backup tools provide a safeguard for SMEs to mitigate the risks of data loss, whether due to human error, cyber-attacks, or technical malfunctions (Thomson et al., 2022). Backup tools also allow SMEs to keep a record of important documents, data and files, thus, enabling them to recover from potential data loss or disruption. This may help to increase the SMEs resilience and improve the overall continuity of their business operations.

Finally, platforms and tools that have been created to support online payment systems were in the focus of Topic 15, "Payment systems" (keyness = 78.49; *p-value* = 0.006). The lack of proper functioning and digital compatibility can cause problems within the SME industry. The shift towards digital payments and online transactions during the pandemic has highlighted the importance of having secure and reliable payment systems in place. SMEs that are not equipped with the necessary tools and platforms to facilitate online payments may face challenges in terms of lost revenue and customer satisfaction (Wu et al., 2022). Additionally, the implementation of digital payment systems may require significant investment and technical expertise, making it a challenge for some SMEs to adopt and effectively utilize these systems (Caloghirou et al., 2022). However, for SMEs who are able to effectively implement digital payment systems, it can provide a significant competitive advantage and a new revenue stream (Bustos-Contell et al., 2021; Dvorsky et al., 2021).

5.1 Theoretical implications

The theoretical implications of the present study are diverse, as our research was conducted using a sample from social networks. In the coronavirus pandemic, social networks have been the main channel of communication for SMEs. This fact has allowed us to analyze relevant data with data mining techniques that identify and respond to the research objective. Our results provide meaningful insights about business adaptability and flexibility of SMEs' business models. Specifically, the research findings indicate that the main innovations and new technologies adopted by SMEs during the COVID-19 pandemic were largely perceived positively, with a focus on tools and solutions that support SMEs' ability to adapt and respond to the changes brought about by the pandemic. Additionally, the findings highlight the importance of cybersecurity and issues with government support.

Also, based on the results presented, it could be argued that the COVID-19 pandemic has accelerated the adoption of new technologies and innovations among SMEs, as they seek to adapt their business models to the changing landscape. One possible implication of this finding is that in times of crisis, there is an increased need for businesses to be agile and adaptable in order to survive. Thus, future research could focus on the specific strategies and technologies that SMEs use in times of crisis, such as free support and webinars, to better understand how they are able to adapt and overcome challenges.

Another implication of the results is that SMEs have increasingly turned to social media and e-commerce as a way to reach customers and generate revenue. This finding highlights the importance of digital channels for SMEs in today's market, and future research could explore how businesses can effectively leverage these channels to maximize their growth and success. The findings also suggest that SMEs have sought out tools and software to improve their efficiency and time management, such as time optimizer and business solutions tools. This implies that there is a growing need for tools that help businesses improve their processes and workflow.

Additionaly, this work extends a research line deploying social media analytics to understand what individual and organizations say about a specific phenomenon such as surveillance technologies (Ribeiro-Navarrete et al., 2021), mechanical artificial intelligence (Mariani & Borghi, 2021a), sustainability issues (Mariani & Borghi, 2020), robots (Borghi and Mariani, 2021) and remote working (Saura et al., 2022a, 2022b, 2022c). This work also theoretically extends the existing body of literature that has used online discourse such as and Mariani and Borghi (2022) to make sense of several phenomena. The current research not only broadens the academic literature on analytics, but also postulates that SMEs will progressively show greater interest in cognitive analytics. This shift is anticipated given that cognitive analytics has been crucial in propelling the growth of large corporations, in conjunction with predictive analytics (Hair et al., 2022; Mariani & Wirtz, 2023; Rocchetta & Mina, 2019).

Future research could examine how these tools are used in different industries and what impact they have on the performance of SMEs. Lastly, the results show that cybersecurity problems and negative sentiment towards government support were also prevalent among SMEs. This implies that SMEs are facing significant challenges in terms of both technology and policy, and future research could explore how businesses are addressing these challenges and what strategies they are using to mitigate risks and build resilience. Finally, the identified topics and their sentiments can be used in statistical models as constructors or variables to measure the relationship between them and to elaborate new scientific contributions.

5.2 Practical implications

With regard to practical applications, the results of the present study SMEs provide a better understanding of the main applications of innovation and technology by SMEs. These uses may help other SMEs develop activities in the same sector. Our results can be used as a guide for the application or adoption of strategies focused on both technological innovation and the acquisition of new tools that would allow businesses to better adapt to the current COVID-19 pandemic or to future pandemics. From prespective, SMEs can promote the adoption of webinar and time optimization tools, as these have been shown to have a positive impact. These tools can assist SMEs in remaining competitive and continuing to operate during future pandemics through the implementation of remote work modes. Additionally, SMEs can use this study to carefully consider the impact of government aid on decision-making. Given that the results indicate a negative impact, SMEs should be aware of how government aid can affect their ability to operate effectively. Furthermore, it is important to highlight that SMEs should capitalize on opportunities that arise through the utilization of social media and e-commerce, as these tools can provide opportunities for expanding reach and diversifying revenue sources.

6 Conclusions

Based on the data analysis, we identified a total of 15 topics—7 positive, 4 negative, and 4 neutral. The positive topics included "Software for process support", "Time optimizer and efficiency", "Back-up tools", "Business solutions tools", "Advisor tools", "Webinar tools" and "Free support against COVID-19". Furthermore, negative topics included "Customer solutions in the cloud", "Cybersecurity problems", "Government support", and "Payment systems". Finally, the identified neutral topics were "Social media and e-commerce", "Specialized startups software", "Big Data analysis tools", and "CRMs and finance". The different topics were presented, discussed and linked to the main contributions to the literature identified to date. In addition, additional insights were obtained to understand the use that SMEs have made of the adapted technologies. Likewise, the research question presented (*What are the primary innovations and technologies adopted by SMEs during the COVID-19 pandemic to adapt their business models based on UGC analysis on Twitter?*) has been answered and discussed through the analysis of the results presented. In this way, the present study has explored the impact of new technologies on SMEs in the context of the COVID-19 pandemic.

The results suggest that SMEs have used a variety of tools and strategies to adapt to the changing market conditions brought on by the pandemic, including free support against COVID-19, webinar tools, social media and e-commerce, time optimizer and efficiency, and business solution tools. The negative sentiment associated with government support highlights the challenges faced by SMEs in accessing adequate support during the pandemic. Additionally, it was identified the use of advisors tools, payment systems and cybersecurity problems These findings contribute to the broader understanding of the challenges and opportunities facing SMEs in the current environment, and provide insights for policymakers and business leaders to consider when developing support and intervention strategies to help SMEs thrive during and post-pandemic. It is recommended that future studies should also be directed to examine the long-term effects of these technologies and strategies on SMEs, as well as the sustainability of SMEs in the new digital era.

Also, the findings suggest that SMEs have been proactive in adopting new technologies and strategies to adapt to the changing environment brought on by the pandemic. The use of digital platforms and tools, such as webinars, social media, and e-commerce solutions, have been essential for SMEs to continue to operate and reach customers in a digital environment. Additionally, SMEs have also adopted tools and strategies focused on efficiency and optimization, as well as business solutions and advisor tools, to stay competitive and meet the changing needs of their customers. It is important to note that while the results of this study provide valuable insights, further research is needed to gain a deeper understanding of the long-term impact of the pandemic on SMEs and their adoption of new technologies. Additionally, future studies should also focus on examining specific sectors and industries to gain a more detailed understanding of how SMEs are adapting in different contexts.

Concurrently, it is anticipated that SMEs will initiate the adoption of technologies (Audretsch et al., 2022a, 2022b) currently being utilized by larger firms, such as cognitive analytics based on AI over traditional predictive analytics (Hair et al., 2022; Mariani & Wirtz, 2023). There is also an inclination towards the application of generative AI, as opposed to less advanced AI forms (Dwivedi et al., 2023; Mariani et al., 2023). Overall, the present study highlights the need for continued support and resources for SMEs as they navigate the challenging business environment brought on by the COVID-19 pandemic. Finally, limitations of the present study are related to the size of the sample, the content of the analyzed, and the machine learning analysis methods used to obtain the research results.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. The authors did not receive support from any organization for the submitted work.

References

Ahmad, M., Aftab, S., Muhammad, S. S., & Ahmad, S. (2017). Machine learning techniques for sentiment analysis: A review. *International Journal of Multidisciplinary Sciences and Engineering, 8*(3), 27.

Akpan, I. J., Udoh, E. A. P., & Adebisi, B. (2020). Small business awareness and adoption of state-of-the-art technologies in emerging and developing markets, and lessons from the COVID-19 pandemic. *Journal of Small Business & Entrepreneurship*. https://doi.org/10.1080/08276331.2020.1820185

Anand, A., Muskat, B., Creed, A., Zutshi, A., & Csepregi, A. (2021). Knowledge sharing, knowledge transfer and SMEs: Evolution, antecedents, outcomes and directions. *Personnel Review, 50*(9), 1873–1893. https://doi.org/10.1108/PR-05-2020-0372

Ante, L. (2023). How Elon Musk's twitter activity moves cryptocurrency markets. *Technological Forecasting and Social Change, 186*, 122112.

Atanassova, I., & Clark, L. (2015). Social media practices in SME marketing activities: A theoretical framework and research agenda. *Journal of Customer Behaviour, 14*(2), 163–183. https://doi.org/10.1362/147539215X14373846805824

Audretsch, D. B., Belitski, M., Caiazza, R., Günther, C., & Menter, M. (2021a). From latent to emergent entrepreneurship: The importance of context. *Technological Forecasting and Social Change, 175*, 121356.

Audretsch, D. B., Lehmann, E. E., Menter, M., & Wirsching, K. (2021b). Intrapreneurship and absorptive capacities: The dynamic effect of labor mobility. *Technovation, 99*, 102129.

Audretsch, D. B., Belitski, M., & Guerrero, M. (2022a). The dynamic contribution of innovation ecosystems to schumpeterian firms: A multi-level analysis. *Journal of Business Research, 144*, 975–986. https://doi.org/10.1016/j.jbusres.2022.02.037

Audretsch, D. B., Belitski, M., Caiazza, R., Guenther, C., & Menter, M. (2022b). Technology adoption over the stages of entrepreneurship. *International Journal of Entrepreneurial Venturing, 14*(4/5), 379.

Ayodele, T. O. (2010). Types of machine learning algorithms. *New Advances in Machine Learning, 3*, 19–48. https://doi.org/10.5772/9385

Banker, S., & Patel, R. (2016). A brief review of sentiment analysis methods. *International Journal of Information Sciences and Techniques (IJIST), 6*(1/2), 89–95.

Barbier, M., Bompart, M., Garandel-Batifol, V., & Mogoutov, A. (2012). Textual analysis and scientometric mapping of the dynamic knowledge in and around the IFSA community. In *Farming Systems Research into the 21st century: The new dynamic* (pp. 73–94). Springer, Dordrecht.

Barbosa, B., Saura, J. R., & Bennett, D. (2022). How do entrepreneurs perform digital marketing across the customer journey? A review and discussion of the main uses. *Journal of Technology Transfer*. https://doi.org/10.1007/s10961-022-09978-2

Belitski, M., & Rejeb, N. (2022). Does open customer innovation model hold for family firms? *Journal of Business Research, 145*, 334–346.

Bender, E. M., & Friedman, B. (2018). Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Transactions of the Association for Computational Linguistics, 6*, 587–604.

Berger, J., Humphreys, A., Ludwig, S., Moe, W. W., Netzer, O., & Schweidel, D. A. (2020). Uniting the tribes: Using text for marketing insight. *Journal of Marketing, 84*(1), 1–25.

Bhavsar, H., & Manglani, R. (2019). Sentiment analysis of Twitter data using Python. *International Research Journal of Engineering and Technology (IRJET), 6*(3), 510–527.

Biber, D. (2004). If you look at …: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371–405. https://doi.org/10.1093/applin/25.3.371

Bigne, E., Simonetti, A., Ruiz, C., & Kakaria, S. (2021). How online advertising competes with user-generated content in TripAdvisor. A neuroscientific approach. *Journal of Business Research, 123*, 279–288.

Borghi, M., & Mariani, M. M. (2021). Service robots in online reviews: Online robotic discourse. *Annals of Tourism Research, 87*, 103036.

Borghi, M., & Mariani, M. M. (2022). The role of emotions in the consumer meaning-making of interactions with social robots. *Technological Forecasting and Social Change, 182*, 121844.

Bui, Q. T., & Lo, F. Y. (2022). Technology Multinational Enterprises from Emerging Markets: Competitive Interplay of International Entry Timing Decisions. *Sustainable Technology and Entrepreneurship, 1*(3), 100019. https://doi.org/10.1016/j.stae.2022.100019

Bustos-Contell, E., Climent-Serrano, S., & Labatut-Serer, G. (2021). A fuzzy-set qualitative comparative analysis model to predict bank bailouts: A study of the Spanish financial system. *Economic Research-Ekonomska Istrazivanja, 34*(1), 2555–2571. https://doi.org/10.1080/1331677X.2020.1833746

Caballero-Morales, S. O. (2021). Innovation as recovery strategy for SMEs in emerging economies during the COVID-19 pandemic. *Research in International Business and Finance, 57*, 101396.

Cai, M. (2021). Natural language processing for urban research: A systematic review. *Heliyon, 7*(3), e06322.

Caloghirou, Y., Giotopoulos, I., Kontolaimou, A., & Tsakanikas, A. (2022). Inside the black box of high-growth firms in a crisis-hit economy: Corporate strategy, employee human capital and R&D capabilities. *International Entrepreneurship and Management Journal, 18*, 1319–1345. https://doi.org/10.1007/s11365-020-00674-x

Cao, J., Zeng, K., Wang, H., Cheng, J., Qiao, F., Wen, D., & Gao, Y. (2013). Web-based traffic sentiment analysis: Methods and applications. *IEEE Transactions on Intelligent Transportation Systems, 15*(2), 844–853.

Carley, K. (1994). Extracting culture through textual analysis. *Poetics, 22*(4), 291–312.

Ceci, F., & Iubatti, D. (2012). Personal relationships and innovation diffusion in SME networks: A content analysis approach. *Research Policy, 41*(3), 565–579.

Chae, B. K. (2015). Insights from hashtag# supplychain and Twitter analytics: Considering Twitter and Twitter data for supply chain practice and research. *International Journal of Production Economics, 165*, 247–259.

Chaithanapat, P., Punnakitikashem, P., Oo, K. K., Chi, N., & Rakthin, S. (2022). Relationships among knowledge-oriented leadership, customer knowledge management, innovation quality and firm performance in SMEs. *Journal of Innovation & Knowledge, 7*(1), 100162. https://doi.org/10.1016/j.jik.2022.100162

Chang, Y. Y., Hughes, M., & Hotho, S. (2011). Internal and external antecedents of SMEs' innovation ambidexterity outcomes. *Management Decision., 49*(10), 1658–1676.

Chopra, M., Singh, S. K., Gupta, A., Aggarwal, K., Gupta, B. B., & Colace, F. (2022). Analysis & prognosis of sustainable development goals using big data-based approach during COVID-19 pandemic. *Sustainable Technology and Entrepreneurship, 1*(2), 100012. https://doi.org/10.1016/j.stae.2022.100012

Chowdhury, G. G. (2003). Natural language processing. *Annual Review of Information Science and Technology, 37*(1), 51–89.

Chung, H. M., & Gray, P. (1999). Data mining. *Journal of Management Information Systems, 16*(1), 11–16. https://doi.org/10.1080/07421222.1999.11518231

Cios, K. J., Pedrycz, W., & Swiniarski, R. W. (2012). *Data mining methods for knowledge discovery*. Springer.

Cui, D., & Curry, D. (2005). Prediction in marketing using the support vector machine. *Marketing Science, 24*(4), 595–615.

Datta, D., Varma, S., & Singh, S. K. (2017). Multimodal retrieval using mutual information based textual query reformulation. *Expert Systems with Applications, 68*, 81–92.

Daugherty, T., Eastin, M. S., & Bright, L. (2008). Exploring consumer motivations for creating user-generated content. *Journal of Interactive Advertising, 8*(2), 16–25.

Duran, N. D., Hall, C., McCarthy, P. M., & McNamara, D. S. (2010). The linguistic correlates of conversational deception: Comparing natural language processing technologies. *Applied Psycholinguistics, 31*(3), 439–462.

Dvorsky, J., Belas, J., Gavurova, B., & Brabenec, T. (2021). Business risk management in the context of small and medium-sized enterprises. *Economic Research-Ekonomska Istrazivanja, 34*(1), 1690–1708. https://doi.org/10.1080/1331677X.2020.1844588

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., & Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research practice and policy. *International Journal of Information Management, 71*, 102642.

Fayyad, U., & Uthurusamy, R. (2002). Evolving data into mining solutions for insights. *Communications of the ACM, 45*(8), 28–31.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine, 17*(3), 37–37.

Fitriasari, F. (2020). How do small and medium enterprise (SME) survive the COVID-19 outbreak? *Jurnal Inovasi Ekonomi*. https://doi.org/10.22219/jiko.v5i02.11838

Ghiassi, M., Skinner, J., & Zimbra, D. (2013). Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with Applications, 40*(16), 6266–6282. https://doi.org/10.1016/j.eswa.2013.05.057

González-Padilla, P., Navalpotro, F. D., & Saura, J. R. (2023). Managing entrepreneurs' behavior personalities in digital environments: A review. *International Entrepreneurship and Management Journal*. https://doi.org/10.1007/s11365-022-00823-4

Grover, P., Kar, A. K., & Davies, G. (2018). "Technology enabled Health"–Insights from twitter analytics with a socio-technical perspective. *International Journal of Information Management, 43*, 85–97. https://doi.org/10.1016/j.ijinfomgt.2018.07.003

Grover, P., Kar, A. K., Janssen, M., & Ilavarasan, P. V. (2019). Perceived usefulness, ease of use and user acceptance of blockchain technology for digital transactions–insights from user-generated content on Twitter. *Enterprise Information Systems, 13*(6), 771–800.

Guo, H., Guo, A., & Ma, H. (2022). Inside the black box: How business model innovation contributes to digital start-up performance. *Journal of Innovation & Knowledge, 7*(2), 100188. https://doi.org/10.1016/j.jik.2022.100188

Haddi, E., Liu, X., & Shi, Y. (2013). The role of text pre-processing in sentiment analysis. *Procedia Computer Science, 17*, 26–32.

Hair, J. F., Harrison, D. E., & Ajjan, H. (2022). *Essentials of Marketing Analytics*. McGraw Hill.

Haque, M. G., Munawaroh, M., & Sunarsi, D. (2020). Analysis of SMEs culinary marketing strategy during Covid 19 pancemic: A study at "Sate Bebek Cilegon" Resto in Cilegon, Banten. *International Journal of Education, Information Technology, and Others, 3*(2), 447–451.

Hardeniya, N., Perkins, J., Chopra, D., Joshi, N., & Mathur, I. (2016). *Natural language processing: python and NLTK*. Packt Publishing Ltd.

Hasan, A., Moin, S., Karim, A., & Shamshirband, S. (2018). Machine learning-based sentiment analysis for twitter accounts. *Mathematical and Computational Applications, 23*(1), 11.

Hassan, N., Gomaa, W., Khoriba, G., & Haggag, M. (2020). Credibility detection in twitter using word n-gram analysis and supervised machine learning techniques. *International Journal of Intelligent and Engineering and Systems, 13*, 291–300.

Heredia, J., Castillo-Vergara, M., Geldes, C., Carbajal Gamarra, F. M., Flores, A., & Heredia, W. (2022). How do digital capabilities affect firm performance? The mediating role of technological capabilities in the "new normal." *Journal of Innovation & Knowledge, 7*(2), 100171. https://doi.org/10.1016/j.jik.2022.100171

Hilmersson, F. P., & Hilmersson, M. (2021). Networking to accelerate the pace of SME innovations. *Journal of Innovation & Knowledge, 6*(1), 43–49. https://doi.org/10.1016/j.jik.2020.10.001

Hiremath, B. N., & Patil, M. M. (2020). Enhancing optimized personalized therapy in clinical decision support system using natural language processing. *Journal of King Saud University-Computer and Information Sciences., 34*(6), 2840–2848.

Hutchinson, V., & Quintas, P. (2008). Do SMEs do knowledge management? Or simply manage what they know? *International Small Business Journal, 26*(2), 131–154.

Ibáñez, M. J., Guerrero, M., Yáñez-Valdés, C., & Barros-Celume, S. (2021). Digital social entrepreneurship: the N-Helix response to stakeholders' COVID-19 needs. *The Journal of Technology Transfer*. https://doi.org/10.1007/s10961-021-09855-4

Isoaho, K., Gritsenko, D., & Mäkelä, E. (2021). Topic modeling and text analysis for qualitative policy research. *Policy Studies Journal, 49*(1), 300–324.

Juergensen, J., Guimón, J., & Narula, R. (2020). European SMEs amidst the COVID-19 crisis: Assessing impact and policy responses. *Journal of Industrial and Business Economics, 47*(3), 499–510.

Kamal, M. M. (2020). The triple-edged sword of COVID-19: Understanding the use of digital technologies and the impact of productive, disruptive, and destructive nature of the pandemic. *Information Systems Management, 37*(4), 310–317.

Kar, A. K. (2021). What affects usage satisfaction in mobile payments? Modelling user generated content to develop the "digital service usage satisfaction model." *Information Systems Frontiers, 23*(5), 1341–1361.

Kaur, C., & Sharma, A. (2020). Twitter Sentiment Analysis on Coronavirus using Textblob (No. 2974). EasyChair.

Keen, A. (2011). *The Cult of the Amateur: How blogs, MySpace, YouTube and the rest of today's user-generated media are killing our culture and economy*. Hachette UK.

Khatami, F., Scuotto, V., Krueger, N., et al. (2022). The influence of the entrepreneurial ecosystem model on sustainable innovation from a macro-level lens. *The International Entrepreneurship and Management Journal, 18*, 1419–1451. https://doi.org/10.1007/s11365-021-00788-w

Kim, H. S., & Sohn, S. Y. (2010). Support vector machines for default prediction of SMEs based on technology credit. *European Journal of Operational Research, 201*(3), 838–846.

Kobrinskii, B. A., & Blagosklonov, N. A. (2018). Hybrid approach to knowledge extraction: Textual analysis and evaluations of experts. *Open Semantic Technologies for Intelligent Systems, 2*(8), 191–195.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications.

Krishnapuram, R., & Keller, J. M. (1996). The possibilistic c-means algorithm: Insights and recommendations. *IEEE Transactions on Fuzzy Systems, 4*(3), 385–393.

Kristinae, V., Wardana, I., Giantari, I. G. A. K., & Rahyuda, A. (2020). The role of powerful business strategy on value innovation capabilities to improve marketing performance during the COVID-19 pandemic. *Uncertain Supply Chain Management, 8*(4), 675–684.

Kumar, R., Singh, R. K., & Dwivedi, Y. K. (2020). Application of industry 4.0 technologies in SMEs for ethical and sustainable operations: Analysis of challenges. *Journal of Cleaner Production, 275*, 124063.

Kumar, A., & Ayedee, D. (2021). Technology Adoption: A Solution for SMEs to Overcome Problems during COVID-19. *Forthcoming, Academy of Marketing Studies Journal*, *25*(1)

Lacárcel, F. J. S. (2022). Main uses of artificial intelligence in digital marketing strategies linked to tourism. *Journal of Tourism, Sustainability and Well-Being, 10*(3), 215–226. https://doi.org/10.34623/mppf-r253

Lacarcel, F. J., & Huete, R. (2023). Digital communication strategies used by private companies, entrepreneurs, and public entities to attract long-stay tourists: a review. *International Entrepreneurship and Management Journal*. https://doi.org/10.1007/s11365-023-00843-8

Laguía, A., & Moriano, J. A. (2021). Perceived representation of entrepreneurship in the mass media and entrepreneurial intention. *The International Entrepreneurship and Management Journal, 17*, 401–421. https://doi.org/10.1007/s11365-019-00609-1

Lai, L. S., & To, W. M. (2015). Content analysis of social media: A grounded theory approach. *Journal of Electronic Commerce Research, 16*(2), 138.

Laksono, R. A., Sungkono, K. R., Sarno, R., & Wahyuni, C. S. (2019, July). Sentiment analysis of restaurant customer reviews on TripAdvisor using Naïve Bayes. In *2019 12th International Conference on Information & Communication Technology and System (ICTS)* (pp. 49–54). IEEE.

Lee, D., & Lee, J. (2007). Domain described support vector classifier for multi-classification problems. *Pattern Recognition, 40*(1), 41–51.

Lipizzi, C., Iandoli, L., & Marquez, J. E. R. (2015). Extracting and evaluating conversational patterns in social media: A socio-semantic analysis of customers' reactions to the launch of new products using Twitter streams. *International Journal of Information Management, 35*(4), 490–503.

Liu, X. (2020). Analyzing the impact of user-generated content on B2B Firms' stock performance: Big data analysis with machine learning methods. *Industrial Marketing Management, 86*, 30–39.

Liu, X., Burns, A. C., & Hou, Y. (2017). An investigation of brand-related user-generated content on Twitter. *Journal of Advertising, 46*(2), 236–247.

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., & Adam, S. (2018). Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures, 12*(2–3), 93–118.

March-Chordà, I., Adame-Sánchez, C., & Yagüe-Perales, R. M. (2021). Key locational factors for immigrant entrepreneurs in top entrepreneurial ecosystems. *The International Entrepreneurship and Management Journal, 17*, 1049–1066. https://doi.org/10.1007/s11365-021-00746-6

Mariani, M., & Baggio, R. (2022). Big data and analytics in hospitality and tourism: A systematic literature review. *International Journal of Contemporary Hospitality Management, 34*(1), 231–278. https://doi.org/10.1108/IJCHM-03-2021-0301

Mariani, M., & Borghi, M. (2020). Environmental discourse in hotel online reviews: A big data analysis. *Journal of Sustainable Tourism, 29*(5), 829–848.

Mariani, M., & Borghi, M. (2021a). Customers' evaluation of mechanical artificial intelligence in hospitality services: A study using online reviews analytics. *International Journal of Contemporary Hospitality Management., 33*(11), 3956–3976.

Mariani, M., & Borghi, M. (2021b). Are environmental-related online reviews more helpful? A big data analytics approach. *International Journal of Contemporary Hospitality Management., 33*(6), 2065–2090.

Mariani, M., & Borghi, M. (2022). Exploring environmental concerns on digital platforms through big data: the effect of online consumers' environmental discourse on online review ratings. *Journal of Sustainable Tourism*. https://doi.org/10.1080/09669582.2022.2033982

Mariani, M. M., & Matarazzo, M. (2021). Does cultural distance affect online review ratings? Measuring international customers' satisfaction with services leveraging digital platforms and big data. *Journal of Management and Governance, 25*(4), 1057–1078.

Mariani, M., & Wirtz, J. (2023). A critical reflection on analytics and artificial intelligence based analytics in hospitality and tourism management research. *International Journal of Contemporary Hospitality Management*. https://doi.org/10.1108/IJCHM-08-2022-1006

Mariani, M. M., Machado, I., Magrelli, V., & Dwivedi, Y. K. (2022). Artificial intelligence in innovation research: A systematic review, conceptual framework, and future research directions. *Technovation, 122*, 102623.

Marine-Roig, E., & Clave, S. A. (2015). A method for analysing large-scale UGC data for tourism: Application to the case of Catalonia. In *Information and communication technologies in tourism 2015* (pp. 3–17). Springer, Cham.

Martín, J. M. M., & Fernández, J. A. S. (2022a). The effects of technological improvements in the train network on tourism sustainability. An approach focused on seasonality. *Sustainable Technology and Entrepreneurship, 1*(1), 100005. https://doi.org/10.1016/j.stae.2022.100005

Martínez, J. M. G., Carracedo, P., Comas, D. G., & Siemens, C. H. (2022). An analysis of the blockchain and COVID-19 research landscape using a bibliometric study. *Sustainable Technology and Entrepreneurship, 1*(1), 100006. https://doi.org/10.1016/j.stae.2022.100006

McEnery, T., & Hardie, A. (2013). The history of corpus linguistics. *The Oxford Handbook of the History of Linguistics, 727*, 745. https://doi.org/10.1093/oxfordhb/9780199585847.013.0034

Menter, M. (2022). Entrepreneurship and Economic Resilience in Times of Crisis: Insights from the COVID-19 Pandemic. In D. B. Audretsch & I. A. M. Kunadt (Eds.), *The COVID-19 Crisis and Entrepreneurship: Perspectives and Experiences of Researchers, Thought Leaders, and Policymakers* (pp. 97–104). Springer International Publishing.

Metallo, C., Agrifoglio, R., Briganti, P., Mercurio, L., & Ferrara, M. (2021). Entrepreneurial behaviour and new venture creation: The psychoanalytic perspective. *Journal of Innovation & Knowledge, 6*(1), 35–42. https://doi.org/10.1016/j.jik.2020.02.001

Meyners, M., Castura, J. C., & Carr, B. T. (2013). Existing and new approaches for the analysis of CATA data. *Food Quality and Preference, 30*(2), 309–319.

Morstatter, F., Kumar, S., Liu, H., & Maciejewski, R. (2013, August). Understanding twitter data with tweetxplorer. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 1482–1485).

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: An introduction. *Journal of the American Medical Informatics Association, 18*(5), 544–551.

Nemes, L., & Kiss, A. (2021). Social media sentiment analysis based on COVID-19. *Journal of Information and Telecommunication, 5*(1), 1–15.

Nurunnabi, M. (2020). Recovery planning and resilience of SMEs during the COVID-19: experience from Saudi Arabia. *Journal of Accounting & Organizational Change., 16*(4), 643–653.

O'Hern, M. S., & Kahle, L. R. (2013). The empowered customer: User-generated content and the future of marketing. *Global Economics and Management Review, 18*(1), 22–30.

O'Kane, C., Cunningham, J. A., Menter, M., & Walton, S. (2021). The brokering role of technology transfer offices within entrepreneurial ecosystems: An investigation of macro–meso–micro factors. *The Journal of Technology Transfer, 46*(6), 1814–1844.

Onan, A., Korukoglu, S., & Bulut, H. (2016). LDA-based topic modelling in text sentiment classification: An empirical analysis. *Int. J. Comput. Linguistics Appl., 7*(1), 101–119.

Ortigueira-Sánchez, L. C., Welsh, D. H., & Stein, W. C. (2022). Innovation drivers for export performance. *Sustainable Technology and Entrepreneurship, 1*(2), 100013. https://doi.org/10.1016/j.stae.2022.100013

Ostrowski, DA (2015). Using latent dirichlet allocation for topic modelling in twitter. In *Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing (IEEE ICSC 2015)* (pp. 493–497). IEEE.

Palmquist, M. E., Carley, K. M., Dale, T. A., & Roberts, C. W. (1997). Applications of computer-aided text analysis: Analyzing literary and nonliterary texts. In *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts*, 171–189.

Pedauga, L., Sáez, F., & Delgado-Márquez, B. L. (2021). Macroeconomic lockdown and SMEs: the impact of the COVID-19 pandemic in Spain. *Small Business Economics, 58*, 665–688.

Pérez-González, D., Trigueros-Preciado, S., & Popa, S. (2017). Social media technologies' use for the competitive information and knowledge sharing, and its effects on industrial SMEs' innovation. *Information Systems Management, 34*(3), 291–301.

Pollach, I. (2012). Taming textual data: The contribution of corpus linguistics to computer-aided text analysis. *Organizational Research Methods, 15*(2), 263–287.

Radko, N., Belitskci, M., & Kalyuzhnova, Y. (2022). Conceptualising the entrepreneurial university: the stakeholder approach. *The Journal of Technology Transfer, 48*, 955–1044.

Ramage, D., Rosen, E., Chuang, J., Manning, C. D., & McFarland, D. A. (2009, December). Topic modeling for the social sciences. In NIPS 2009 workshop on applications for topic models: text and beyond (Vol. 5, p. 27).

Ray, A., & Bala, P. K. (2021). User generated content for exploring factors affecting intention to use travel and food delivery services. *International Journal of Hospitality Management, 92*, 102730.

Ray, A., Bala, P. K., & Dwivedi, Y. K. (2020). Exploring values affecting e-Learning adoption from the user-generated-content: A consumption-value-theory perspective. *Journal of Strategic Marketing, 29*(5), 430–452.

Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. In The Workshop on Comparing Corpora (pp. 1–6). https://doi.org/10.3115/1117729.1117730.

Ribeiro-Navarrete, S., Saura, J. R., & Palacios-Marqués, D. (2021). Towards a new era of mass data collection: Assessing pandemic surveillance technologies to preserve user privacy. *Technological Forecasting and Social Change, 167*, 120681. https://doi.org/10.1016/j.techfore.2021.120681

Rivna, A., & Gress, M. (2022). Social entrepreneurship as a new business model. The case of refugee integration. *Economic Research-Ekonomska Istrazivanja*. https://doi.org/10.1080/1331677X.2022.2120044

Rocchetta, S., & Mina, A. (2019). Technological coherence and the adaptive resilience of regional economies. *Regional Studies, 53*(10), 1421–1434.

Roper, S., & Turner, J. (2020). R&D and innovation after COVID-19: What can we expect? A review of prior research and data trends after the great financial crisis. *International Small Business Journal, 38*(6), 504–514.

Runco, M. A. (1993). Operant theories of insight, originality, and creativity. *American Behavioral Scientist, 37*(1), 54–67.

Saura, J. R. (2021). Using Data Sciences in Digital Marketing: Framework, methods, and performance metrics. *Journal of Innovation & Knowledge., 6*(2), 92–102. https://doi.org/10.1016/j.jik.2020.08.001

Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2021a). From user-generated data to data-driven innovation: A research agenda to understand user privacy in digital markets. *International Journal of Information Management, 60*, 102331.

Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2021b). Setting B2B digital marketing in artificial intelligence-based CRMs: A review and directions for future research. *Industrial Marketing Management, 98*, 161–178. https://doi.org/10.1016/j.indmarman.2021.08.006

Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2022a). Assessing behavioral data science privacy issues in government artificial intelligence deployment. *Government Information Quarterly, 39*(4), 101679. https://doi.org/10.1016/j.giq.2022.101679

Saura, J. R., Ribeiro-Soriano, D., & Saldaña, P. Z. (2022b). Exploring the challenges of remote work on Twitter users' sentiments: From digital technology development to a post-pandemic era. *Journal of Business Research, 142*, 242–254.

Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2022c). Adopting digital reservation systems to enable circular economy in entrepreneurship. *Management Decision*. https://doi.org/10.1108/MD-02-2022-0190

Saura, J. R., Palacios-Marqués, D., & Ribeiro-Soriano, D. (2023a). Exploring the boundaries of Open Innovation: Evidence from social media mining. *Technovation, 119*, 102447. https://doi.org/10.1016/j.technovation.2021.102447

Saura, J. R., Palacios-Marqués, D., & Ribeiro-Soriano, D. (2023b). Privacy concerns in social media UGC communities: Understanding user behavior sentiments in complex networks. *Information Systems and e-Business Management*. https://doi.org/10.1007/s10257-023-00631-5

Saura, J. R., Ribeiro-Navarrete, S., Palacios-Marqués, D., & Mardani, A. (2023c). Impact of extreme weather in production economics: Extracting evidence from user-generated content. *International Journal of Production Economics, 260*, 108861. https://doi.org/10.1016/j.ijpe.2023.108861

Scuotto, V., Del Giudice, M., & Obi Omeihe, K. (2017). SMEs and mass collaborative knowledge management: Toward understanding the role of social media networks. *Information Systems Management, 34*(3), 280–290.

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge University Press.

Shokhnekh, AV., Mironova, OA., Hanafeev, FF., Kuzmenko, OA., & Shilova, LF. (2019). Indicators of artificial intelligence of financial evaluation of small business investment attractiveness. In *Ubiquitous Computing and the Internet of Things: Prerequisites for the Development of ICT* (pp. 1031–1041). Springer, Cham.

Short, J. C., Broberg, J. C., Cogliser, C. C., & Brigham, K. H. (2010). Construct validation using computer-aided text analysis (CATA) an illustration using entrepreneurial orientation. *Organizational Research Methods, 13*(2), 320–347.

Slatten, L. A., Bendickson, J. S., Diamond, M., & McDowell, W. C. (2021). Staffing of small nonprofit organizations: A model for retaining employees. *Journal of Innovation & Knowledge, 6*(1), 50–57. https://doi.org/10.1016/j.jik.2020.10.003

Smith, A. N., Fischer, E., & Yongjian, C. (2012). How does brand-related user-generated content differ across YouTube, Facebook, and Twitter? *Journal of Interactive Marketing, 26*(2), 102–113.

Sulistyo, H. (2016). Innovation capability of SMEs through entrepreneurship, marketing capability, relational capital and empowerment. *Asia Pacific Management Review, 21*(4), 196–203.

Susarla, A., Oh, J. H., & Tan, Y. (2012). Social networks and the diffusion of user-generated content: Evidence from YouTube. *Information Systems Research, 23*(1), 23–41.

Syaifullah, J., Syaifudin, M., Sukendar, M. U., & Junaedi, J. (2021). Social Media Marketing and Business Performance of MSMEs During the COVID-19 Pandemic. *The Journal of Asian Finance, Economics, and Business, 8*(2), 523–531.

Syriopoulos, K. (2020). The impact of COVID-19 on entrepreneurship and SMEs. *Journal of the International Academy for Case Studies, 26*(2), 1–2.

Tajbakhsh, M. S., & Bagherzadeh, J. (2019). Semantic knowledge LDA with topic vector for recommending hashtags: Twitter use case. *Intelligent Data Analysis, 23*(3), 609–622.

Thomson, L., Kamalaldin, A., Sjödin, D., & Parida, V. (2022). A maturity framework for autonomous solutions in manufacturing firms: The interplay of technology, ecosystem, and business model. *International Entrepreneurship and Management Journal, 18*, 125–152. https://doi.org/10.1007/s11365-020-00717-3

Tirunillai, S., & Tellis, G. J. (2012). Does chatter really matter? Dynamics of user-generated content and stock performance. *Marketing Science, 31*(2), 198–215.

Tufekci, Z. (2014, May). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 8, No. 1).

Viaene, S., Baesens, B., Van Gestel, T., Suykens, J. A. K., Van den Poel, D., Vanthienen, J., & Dedene, G. (2001). Knowledge discovery in a direct marketing case using least squares support vector machines. *International Journal of Intelligent Systems, 16*(9), 1023–1036.

Vivona, R., Demircioglu, M. A., & Audretsch, D. B. (2022b). The costs of collaborative innovation. *The Journal of Technology Transfer, 48*, 873–899.

Wang, L. (Ed.). (2005). *Support vector machines: theory and applications*. Springer. https://doi.org/10.1007/b95439

Wang, Y., Hong, A., Li, X., & Gao, J. (2020). Marketing innovations during a global crisis: A study of China firms' response to COVID-19. *Journal of Business Research, 116*, 214–220.

Warner-Søderholm, G., Bertsch, A., Sawe, E., Lee, D., Wolfe, T., Meyer, J., & Fatilua, U. N. (2018). Who trusts social media? *Computers in Human Behavior, 81*, 303–315.

Wolfe, M. T., Patel, P. C., & Manikas, A. S. (2021). Shock and awe: Loudness and unpredictability in Twitter messages and crowdfunding campaign success. *Journal of Innovation & Knowledge., 6*(4), 246–256. https://doi.org/10.1016/j.jik.2021.06.002

Wu, Y., Zhao, C., & Guo, J. (2022). Mobile payment and subjective well-being in rural china. *Economic Research-Ekonomska Istrazivanja*. https://doi.org/10.1080/1331677X.2022.2097103

Xue, J., Chen, J., Chen, C., Zheng, C., Li, S., & Zhu, T. (2020). Public discourse and sentiment during the COVID 19 pandemic: Using Latent Dirichlet Allocation for topic modeling on Twitter. *PLoS ONE, 15*(9), e0239441.

Yang, M. C., & Rim, H. C. (2014). Identifying interesting Twitter contents using topical analysis. *Expert Systems with Applications, 41*(9), 4330–4336.

Yi, H., Meng, X., Linghu, Y., & Zhang, Z. (2022). Can financial capability improve entrepreneurial performance? evidence from rural china. *Economic Research-Ekonomska Istrazivanja*. https://doi.org/10.1080/1331677X.2022.2091631

Yu, D., Xu, D., Wang, D., & Ni, Z. (2019). Hierarchical topic modeling of Twitter data for online analytical processing. *IEEE Access, 7*, 12373–12385.

Zeng, S. X., Xie, X. M., & Tam, C. M. (2010). Relationship between cooperation networks and innovation performance of SMEs. *Technovation, 30*(3), 181–194.

Zhao, X., & Jiang, J. (2011). An empirical comparison of topics in twitter and traditional media. *Singapore Management University School of Information Systems Technical paper series. Retrieved November*, *10*, 2011.

Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E. P., Yan, H., & Li, X. (2011, April). Comparing twitter and traditional media using topic models. In *European conference on information retrieval* (pp. 338–349). Springer, Berlin, Heidelberg.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Jose Ramon Saura[1] · **Daniel Palacios-Marqués**[2] · **Domingo Ribeiro-Soriano**[3]

✉ Jose Ramon Saura
joseramon.saura@urjc.es

Daniel Palacios-Marqués
dapamar@doe.upv.es

Domingo Ribeiro-Soriano
domingo.ribeiro@uv.es

1 Rey Juan Carlos University, Madrid, Spain

2 Universitat Politècnica de València, Valencia, Spain

3 Universitat de Valencia, Valencia, Spain