

Using Social Media to Extract Indirect Experiential Knowledge for the Internationalization of an Entrepreneurial Firm

Johannes Järviemi, Heli Hallikainen and Tommi Laukkanen

University of Eastern Finland, Finland

heli.hallikainen@uef.fi

10.34190/EJKM.18.01.004

Abstract: Social media platforms have become popular and widely used, and they contain large amounts of knowledge hidden in the user-generated content. This study, hence, suggests that analyzing the semantic meanings hidden within the user-generated content on social media platforms enables the extraction of knowledge about international markets which has not been previously available, but that can enhance firms' understanding of foreign markets. The study suggests that novel methods of content analytics, including sentiment analysis, can provide entrepreneurial firms with a cost-effective way to acquire knowledge about foreign markets in general and about customers, competitors and other stakeholders operating in the international supply chain. To demonstrate this, we present a case study of a craft brewery and use sentiment analysis to extract knowledge of craft beer related discussions in general and of two brands currently operating in the pre-specified target markets of a craft brewery. Based on an analysis of 20,153 tweets extracted in a two-weeks data collection period from Twitter, the results of the study show that, in general, craft beer related discussions are positively rather than negatively oriented, and hence there may exist potential for the products of the case company. In addition, the results of the study provide insights into two competing brands currently operating in foreign markets. When comparing the tweets of the two competing brands currently operating in the target markets, the results of the study showed that the tweets related to brand B were more positively oriented and associated more frequently with positive sentiments, in comparison to brand A.

Keywords: Internationalization, Social media, Twitter, Market knowledge, Indirect experiential knowledge

1. Introduction

Being active on various social media platforms seems no longer to be a trend but rather a norm for large firms and multinational enterprises. Social media platforms, i.e. Internet-based channels enabling interaction and self-presentation online (Carr and Heyes, 2015), are being widely used, and a study conducted by the University of Massachusetts Dartmouth reports that 98 percent of Fortune 500 companies actively use LinkedIn, and 91 and 89 percent of the same companies are actively using Twitter and Facebook, respectively (Barnes, Kane and Maloney, 2018). Corporate blogs have also become more prevalent for companies, while there has been an 11 percent increase in blog usage since 2017, and 53 percent of Fortune 500 companies now have corporate blogs (Barnes, Kane and Maloney, 2018). International firms publish a wealth of information and invest large amounts of money in their presence on various social media platforms to create and strengthen their relationships and interaction with customers, build brand awareness and improve their customer engagement, among other expected outcomes (Tsimonis and Dimitriadis, 2014). Customers interact and engage with companies through different social media platforms and add their share to this. Importantly, the data available on social media platforms is widely available to the public. Hence, social media platforms provide a wealth of data, containing hidden knowledge which companies can leverage for their competitive advantage (He, Zha and Li, 2013). Additionally, this knowledge is available not only for the large multinational enterprises but also for smaller entrepreneurial firms. Cesaroni and Consoli (2015) propose that entrepreneur's mentality and attitudes make a difference in the ability to take full advantage of social media platforms.

While the existing literature has established the key role of knowledge in the internationalization process of large firms (e.g. Buckley et al., 2016), very little research exists about how knowledge acquired from social media can support the internationalization plans of small and medium sized entrepreneurial firms. The present study, therefore, demonstrates the use of content analytics together with data extracted from Twitter, and demonstrates the use of such methods to support the internationalization process of an entrepreneurial firm. Following the concept of independent entrepreneurship, entrepreneurial firms in this study refer to independent entrepreneurship, i.e. "the process whereby an individual or a group of individuals, acting independently of any association with an existing organization, create a new organization" (Sharma and Chrisman, 2007, p. 92). With regard to internationalization, international entrepreneurship refers to the

discovery, enactment, evaluation and exploitation of opportunities across national borders to create goods and services (Oviatt and McDougall, 2005).

When starting an international business, data extracted from social media platforms can be useful for discovering new knowledge, such as information about the popularity of brands (De Vries, Gensler, and Leeftang, 2012) and in attempts to understand changing patterns and trends, such as what competitors are doing and how the industry is changing (Dey et al., 2011; He, Zha and Li, 2013). Vasileiadou and Missler-Behr (2011) report that new ventures are keen on utilizing social media particularly for internal and external communication and collaboration purposes. Although social media related studies have become increasingly popular in the past years, studies that attempt to demonstrate how unstructured data on social media can be harnessed to extract business knowledge are still relatively scarce in the scholarly literature (Balducci and Marinova, 2018; Nguyen et al., 2015; Pathak and Pathak-Shelat, 2017). A particular paucity of research exists regarding the implications of social media usage considering the internationalization process of the company, although theoretical research highlights the potential influence of the firm's social media usage on international business performance (Berthon et al., 2012; Maltby, 2012; Okazaki and Taylor, 2013). Indeed, potential exists as Fraccastoro and Gabriellsson (2018) suggest that the usage of social media enables small firms to effectively reach weak network ties, and through such ties firms can enhance the speed of international growth.

In this paper, we propose that unstructured data extracted from Twitter can provide useful market knowledge for entrepreneurial firms and support their strategic plans to internationalize and expand their operations into foreign markets. We demonstrate how sentiment analysis, which is a fairly novel data analysis method, can be used to extract knowledge arising from social media data. In the study we focus specifically on smaller firms which are typically led by entrepreneurs because Fletcher and Harris (2012) highlight that smaller firms may not have relevant experience or useful networks and may therefore need to rely on sources of information rarely recognized in the literature. Hence, the study proposes that publicly available data on social media platforms harnessed using content analytics can help the companies to gain knowledge about their competing companies and brands in foreign markets. Such methods can help the companies to extract insights of customer opinions and sentiments regarding competing products and brands. Analyzing the semantic meanings hidden in online discussions enables the extraction of external indirect experiential market knowledge, defined as "knowledge developed by others but accessible by the firm" (Åkerman, 2015; Fletcher and Harris, 2012) about international markets which has not been previously available.

The present study focuses on initial phases of the internationalization timeline as depicted in the knowledge-based model of internationalization (Mejri and Umemoto, 2010). During the initial phases of internationalization, the role of external market knowledge is particularly relevant as firms typically have very little or no knowledge of the prospective markets (Mejri and Umemoto, 2010). Consequently, this paper makes an effort to describe how firms can extract knowledge about foreign markets from discussions on social media and use this to enhance their understanding about those markets. We demonstrate this by introducing a business case showing how sentiment analysis and publicly available data on Twitter was used to extract knowledge about competitors and competing brands to support a small entrepreneurial brewing company's intention to enter foreign craft beer markets.

The remainder of the paper unfolds as follows. In the next section, we review the theoretical background of the study and explain our reasoning for adopting it as the theoretical framework of this study. This is followed by a description of the research model and research methods. We then present the results of the data analysis, and discuss the theoretical and practical implications of it together with the limitations of the study and our suggestions for future research avenues.

2. Theoretical background

Knowledge plays a pivotal role in a firm's internationalization (Åkerman, 2015; Mejri, Alexander and Tsagdis, 2018; Stoian, Dimitratos and Plakoyiannaki, 2018). Following several prior studies on the role of knowledge as a driver of internationalization (Deligianni, Voudouris and Lioukas, 2015; Gassmann and Keupp, 2007), we ground our study on the knowledge-based view (KBV) (Grant, 1996), which itself is grounded on and extends the resource-based view (RBV) (Wernerfelt, 1984). Specifically, the knowledge-based view emphasizes the role of intangible resources and the optimization of knowledge as a strategic asset and a source of superior firm performance (DeCarolis and Deeds, 1999).

Market knowledge, i.e., information about foreign markets (Mejri and Umemoto, 2010) and knowledge about the local business context (Åkerman, 2015), is crucial when a firm plans to expand its operations into foreign markets (Johanson and Vahlne, 1977). Furthermore, foreign market knowledge may be the driving factor that differentiates traditional models of internationalization from firms that internationalize early (Zhou, 2007). Foreign market knowledge acquired during the initial stages of internationalization is critical for the start of internationalization, particularly prior to gaining experiential knowledge (Mejri and Umemoto, 2010). The difficulties of dealing with local governments and partners, and the challenges of adapting the products and processes to different cultures and consumer preferences increases the risk of making costly errors which may lead to substantial delays, among other struggles to establishing successful operations abroad (Lord and Ranft, 2000). Market knowledge therefore has a key role in internationalization (Zhou, 2007) by reducing uncertainty and risks as firms expand into international markets (Fletcher and Harris, 2012; Freeman et al., 2010).

Eriksson et al. (1997; 2000) suggests that a lack of foreign market knowledge is a significant obstacle to a firm's internationalization. A company operating in international markets should therefore have the capability to acquire, interpret and integrate intelligence and knowledge about trends, customer preferences and environmental changes in order to identify present and future opportunities in those markets (Yeniyurt, Cavusgil and Hult, 2005). However, small firms in particular may not have relevant experience or useful networks in international markets (Fletcher and Harris, 2012), and hence they need to rely on other sources of information. This study suggests that unstructured data obtained from social media and analyzed using content analytics can provide fruitful market knowledge to support the firm's plans to internationalize.

Eriksson et al. (1997) identified three types of foreign market knowledge: foreign institutional knowledge (knowledge of a foreign culture, and its institutions, rules and regulations), foreign business knowledge (knowledge pertaining to customers, competitors and markets conditions), and internationalization knowledge (knowledge in terms of adapting resources and capabilities to engage in international operations). The first two types of knowledge assist firms in identifying opportunities and problems in the foreign market, while the third knowledge type enables the firm to take applicable steps to open up new markets (Zhou, 2007). With regard to knowledge acquisition, De Clercq et al. (2012) derive five distinct ways of acquiring knowledge from the international entrepreneurship literature: experiential learning, vicarious learning, searching, grafting and congenital learning.

Based on prior research on internationalization, Fletcher and Harris (2012) distinguish between internal and external sources of knowledge, and experiential and objective knowledge. Objective knowledge such as market research, government statistics and company reports are relatively easy to obtain, while experiential knowledge is more challenging to acquire and transfer (Eriksson et al., 1997; Fletcher and Harris, 2012; Nonaka, 1994). Hilmersson (2014) further categorizes experiential knowledge types into internationalization knowledge, institutional knowledge, business network knowledge and social network knowledge. Based on the earlier literature, Pellegrino and McNaughton (2017) categorize external sources of knowledge as: 1) alliance partners, venture capital firms and proximal firms, 2) industry organizations, consultants, top management's congenital knowledge, 3) newly recruited managers, partners, suppliers, customers, consultants and industry/government organizations and 4) network relationships. Additionally, internal sources of knowledge include the prior knowledge of the entrepreneur, the top management team and other key members within the firm.

The distinctions between internal vs external knowledge; and experiential vs objective knowledge enables four different sources for knowledge acquisition to be derived (Figure 1) (Åkerman, 2015; Fletcher and Harris, 2012; Huber, 1991). The focus in this research is on external knowledge which is indirect and experiential. This has received notably little attention in the existing body of literature. In the literature, indirect experiential knowledge refers, e.g., to vicarious learning (i.e. imitative learning or learning by observing others) and grafting (i.e., the acquisition of firms / business units and recruiting personnel) (Åkerman, 2015; Casillas, Barbero and Sapienza, 2015; De Clercq et al., 2012; Fletcher and Harris, 2012) and to learning from the mistakes made by others (Kuivalainen, Saarenketo and Puumalainen, 2012). This study proposes that social media can be a new fruitful platform for obtaining indirect experiential market knowledge. Specifically, indirect experiential market knowledge obtained from social media can be useful in the firm's attempts to understand different market conditions; customers, their behavior and preferences; competitors, suppliers and other stakeholders.

	Internal knowledge	External knowledge
Experiential knowledge	<i>Direct experiential knowledge</i> acquired by the firm's own on-going operations.	<i>Indirect experiential knowledge</i> which is developed by others but can be accessed by the firm.
Objective knowledge	<i>Internal information</i> consists of knowledge obtained by the firm and its employees.	<i>External search</i> results in objective knowledge obtained from external sources.

Figure 1: Knowledge acquisition sources in internationalization (Åkerman, 2015; Fletcher and Harris, 2012).

2.1 Knowledge acquisition from social media

Knowledge acquisition refers to “the process by which knowledge is obtained” (Huber, 1991, p. 90) and Nguyen et al. (2015) define knowledge acquisition from social media as the capability to accumulate adequate and critical knowledge arising from social media. Knowledge acquisition hence aims to gain intelligence, i.e. some purposeful meaning from data (Bose, 2008). Fleischer (2008, p. 853) defines intelligence as “a systematic, targeted, timely and ethical effort to collect, synthesize and analyze competition, markets and the external environment in order to produce actionable insights for decision-makers”. Hence, intelligence differs from data and information because it requires some form of analysis to derive meaning from the data (Bose, 2008). For expanding into international markets, knowledge acquisition from social media can be useful for identifying opportunities and threats in the target markets, particularly as social media users tend to be expressive of their opinions and about how they use different products and services (Nguyen et al., 2015).

Market intelligence here is considered an umbrella term, and based on recent studies (Katsikea, Theodosiou and Makri 2019; Schoenherr and Swink, 2015), customer intelligence, competitor intelligence and supply chain intelligence are considered to be subcategories of market intelligence (Figure 2). Market intelligence, in general, refers to the organization-wide gathering, sharing and use of exogenous market factors that affect customer needs and preferences, as well as current and future needs of customers (Kohli and Jaworski, 1990; Haverila and Ashill, 2011). In this study, market intelligence refers to external intelligence obtained from social media and it is collected by a firm about a specific market it wishes to enter. Competing firms typically have access to the same sets of external data, and competitive advantages do not arise per se from the access to such data, but rather in a firm’s ability in using and applying it (Maltz and Kohli, 1996).

Customer intelligence in this study refers to intelligence obtained about customers in international markets. Information about customer opinions and preferences can be obtained by monitoring general discussions and opinions about a specific product / service category (e.g., craft beer in the case example using hashtag #craftbeer) using for instance sentiment analysis and semantic analysis methods.

Competitor intelligence here refers to external intelligence about the firm’s competitors and various text mining methods, such as sentiment analysis and semantic analysis methods that can be used to extract knowledge regarding competitors in international markets. Competitor intelligence can be obtained on two levels: 1) focusing on the content generated by competing firms themselves and 2) focusing on the content generated by customers of the competing firms.

Additionally, *supply chain intelligence* in this study extends the intelligence to various stakeholders operating in the international markets. Utilizing social media data, visualization tools can be used to create network graphs of different operators in international markets. Kim and Hastak (2018) illustrate how social media network analytics were applied to study characteristics of online social networks after a disaster, and similar methods could be used to visualize the links and connections in an international supply chain.

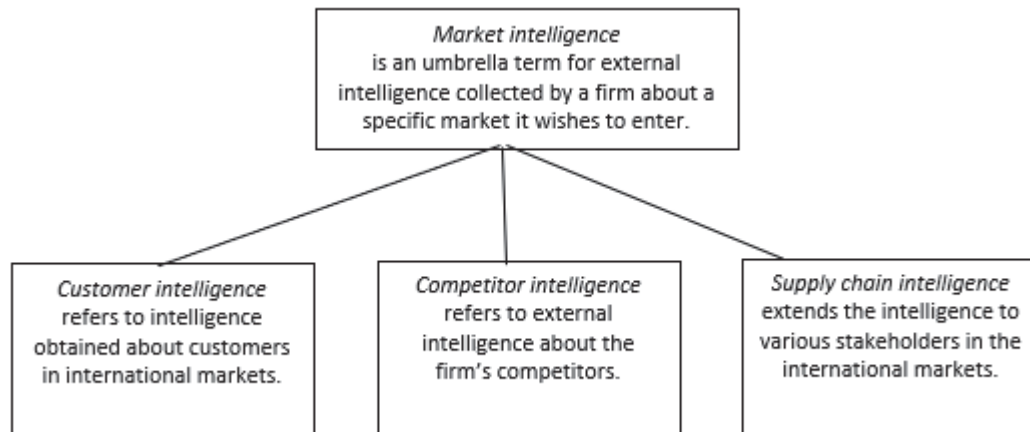


Figure 2: A typology of intelligence extracted from social media.

3. Business case

In this study we conducted a case study to demonstrate how content analytics methods together with unstructured data extracted from Twitter could be used to gain external market knowledge for a craft brewery in the pre-internationalization phase of the internationalization. The case company had approximately 100 employees and at the time of this research, it had operated solely in its home country. The company had a strategic plan to expand into international markets, but it did not have much market knowledge about the prospective markets. The exploratory case hence describes the process of extracting knowledge for the use of the case company and illustrates how content analytics methods and unstructured data obtained from Twitter were harnessed to support the company's plans to internationalize. Specifically, the data analysis of this study is grounded on 20,153 Twitter tweets obtained using DiscoverText and the Gnip-enabled Twitter PowerTrack 2.0 solution. The data gathered was then examined using sentiment analysis methods and three alternative sentiment lexicons: Bing, AFINN and NRC.

3.1 Data collection

Data for the empirical study was collected from Twitter using DiscoverText and the Gnip-enabled Twitter PowerTrack 2.0 solution which enables Twitter tweets to be extracted based on filtering criteria specified by the user. In setting up the extraction criteria, we limited the tweets to a specific country that operated as a pre-defined target market for the company's plans to internationalize. Within a period of two weeks, 20,153 tweets were collected for the purpose of this study. Two different approaches were used to searching through the data for the study. Because the case company operated in the craft beer industry, we first followed general craft beer related content using the term craft beer (referred to as #craftbeer dataset in this study) as a search criterion. Additionally, we focused on two brands known to be operating in the prospective market area to obtain information about the competing brands (referred to as brand A and brand B datasets).

From DiscoverText, the data was extracted in CSV format, and then analyzed using the tidytext package available in the RStudio environment. First, the data was cleaned and prepared using tidy data principles (Wickham, 2014). This phase, referred to as preprocessing (García et al., 2016), is one of the most important phases in the process of knowledge discovery (Han and Kamber, 2001), and this is especially so in this case since the data extracted from Twitter is likely to be imperfect and contain numerous inconsistencies and redundancies. Indeed, previous studies have reported that user-generated content in general tends to be quite informal, and it is characteristics to tweets that users create their own words and slang, spellings, punctuation, abbreviations etcetera (Singh and Kumari, 2016).

3.2 Text mining and sentiment analysis

Text mining methods were used to analyze the cleaned and prepared data. Text mining methods can be used to extract meaningful information from unstructured textual data such as text files, websites, chat messages and emails (Cao, Duan and Gan, 2011; Tang, Pei and Luk, 2014). Text mining and various applications of it that attempt to identify models, trends, patterns and rules within textual data, are being increasingly used to analyze publicly available data obtained from online communities and social media (He, Zha and Li, 2013; Mostafa, 2013).

Sentiment analysis is an approach to text mining, and it is described as “the computational task of automatically detecting and classifying sentiment from text, depending on its polarity or orientation” (Balahur and Jacquet, 2015, p. 428). Sentiment lexicon is a key element of sentiment analysis and it contains a sentiment value describing the polarity of words or sentences (e.g. positive, negative) (Oliveira, Cortez and Areal, 2016). Sentiments represent emotions, and emotions play an important role in marketing decisions (Bougie, Pieters and Zeelenberg, 2003; Gaur, Herjanto and Makkar, 2014). Thus, emotional information obtained from online communities can help to determine the likelihood or success of a particular product or a service (Pathak and Pathak-Shelat, 2017). In this study, three widely used sentiment lexicons were used to analyze the textual data extracted from Twitter. These were the Bing, AFINN and NRC lexicons:

The *Bing lexicon* is a polarity-oriented lexicon and it uses 2,006 positive words and 4,683 negative words to form the results (Bravo-Marquez, Mendoza and Poblete, 2014). The use of the Bing lexicon allows to determine the frequencies of positive and negative words in a dataset, and to analyze how these words fluctuate within a dataset.

The *AFINN lexicon* is based on the Affective Norms for English Words (ANEW) lexicon proposed by Bradley and Lang (2008). Discussion on social media, however, tends to be more informal and thus has a poorer lexical quality compared to many other formats of written text (Rello and Baeza-Yates, 2012). Therefore, efforts were taken by Nielsen (2011) to adapt and fine tune the lexicon specifically for analyzing sentiments found in microblogging. The AFINN lexicon provides affective ratings for a large number of English words based on their valence (Bravo-Marquez, Mendoza and Poblete, 2014) and assigns them a score ranging between -5 and 5.

The *NRC lexicon* was developed by Mohammad and Turney (2013), and it contains over 14,000 English words annotated following Plutchik’s (2002) wheel of emotion. Figure 3 illustrates the text mining process applied in the study.

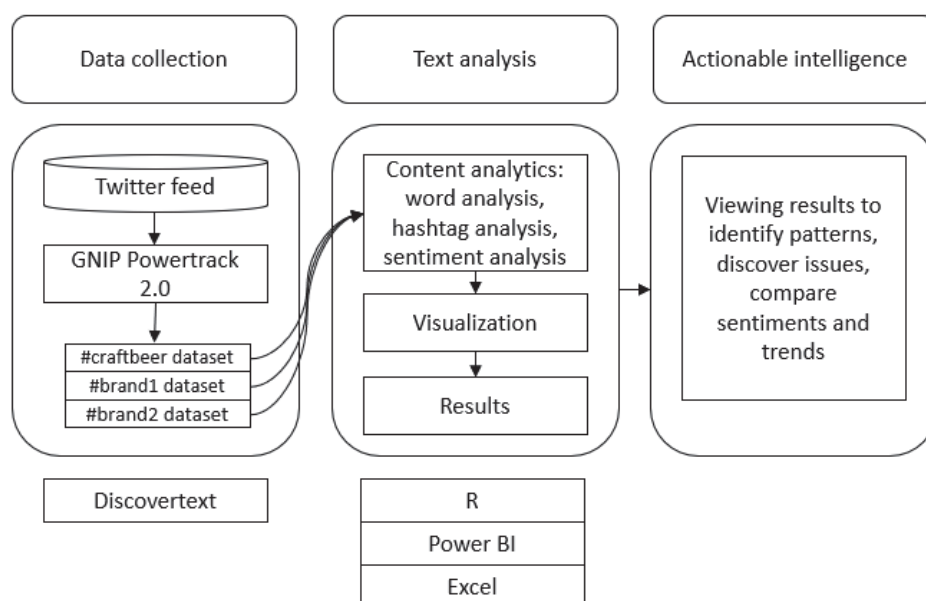


Figure 3: The text mining process used in the study (adapted from He, Zha and Li, 2013).

3.3 Results

The analysis was initiated using the Bing lexicon which provides information about the polarity of the discussion and classifies the words as either positive or negative. The results of the study show that in the target markets, the Twitter discussion about craft beers is mainly positively oriented as the Bing lexicon classifies 69 percent of the discussion in the #craftbeer dataset as positive and 31 percent as negative. A comparison between brands A and B reveals that the tweets relating to brand B are more positively oriented, compared to brand B (Table 1). Table 2 further illustrates a listing of the top words that contributed to the results of the classification in the #craftbeer dataset.

Table 1: The results of applying the Bing lexicon to the dataset.

Sentiment	#craftbeer dataset	Brand A dataset	Brand B dataset
Positive (%)	69	49	68
Negative (%)	31	51	32
Number of Tweets	25,487	16,598	4,262

Table 2: The top 10 positive and negative words using the Bing lexicon.

#	Top 10 positive words			Top 10 negative words		
	Word	N	% of all	Word	N	% of all
1	Love	501	7.54	Pale	199	6.69
2	Happy	337	5.07	Sour	154	5.18
3	Delicious	245	3.69	Limited	91	3.06
4	Fun	194	2.92	Dark	90	3.03
5	Enjoy	189	2.85	Cold	85	2.86
6	Favorite	169	2.54	Miss	80	2.69
7	Awesome	151	2.27	Hard	62	2.09
8	Free	138	2.08	Hazy	58	1.95
9	Amazing	129	1.94	Bad	49	1.65
10	Ready	125	1.88	Drunk	48	1.61

Next, an analysis using the NRC lexicon provides more detailed information about the sentiments present in the Twitter discussions. In all three datasets positive, joy and anticipation are the most frequently present sentiments. However, a comparison between brands A and B show clear differences, and the discussion about brand B is more frequently associated with positive sentiments compared to brand A (Table 3). Additionally, negative sentiments, including, sadness, anger, fear and disgust, are slightly more present in brand A dataset, compared to brand B dataset. Using the NRC lexicon, it would be possible to extract a listing of the top words contributing to each sentiment, similarly to what is described in Table 2 demonstrating the use of the Bing lexicon. For practical reasons we do not report those results in detail, however, such an approach is considered helpful in obtaining a more thorough understanding of the associated sentiments.

Table 3: The results of applying the NRC lexicon to the dataset.

Sentiment	#craftbeer dataset		Brand A dataset		Brand B dataset	
	Words	% of all sentiments	Words	% of all sentiments	Words	% of all sentiments
Positive	11,535	45	3,607	22	1,402	33
Joy	6,488	25	2,682	16	596	14
Anticipation	1,542	6	2,097	13	548	13
Trust	1,006	4	1,758	10	427	10
Negative	971	4	1,724	10	397	9
Surprise	933	4	1,177	7	210	5
Sadness	904	3	1,134	7	195	5
Anger	738	3	966	6	186	4
Fear	722	3	800	5	157	4
Disgust	648	3	653	4	144	3

Finally, an analysis using the AFINN lexicon shows that the AFINN scores are 1.436, 0.234 and 0.681 for the #craftbeer, brand A and brand B datasets, respectively. The AFINN score could range between -5 and 5, and thus the discussion in the #craftbeer dataset is the most positively oriented. These scores support findings of our previous analyses as brand B has a higher score compared to brand A. Figure 4 illustrates an example of AFINN results for the #craftbeer dataset showing top 25 words that contribute to the AFINN score of the dataset.

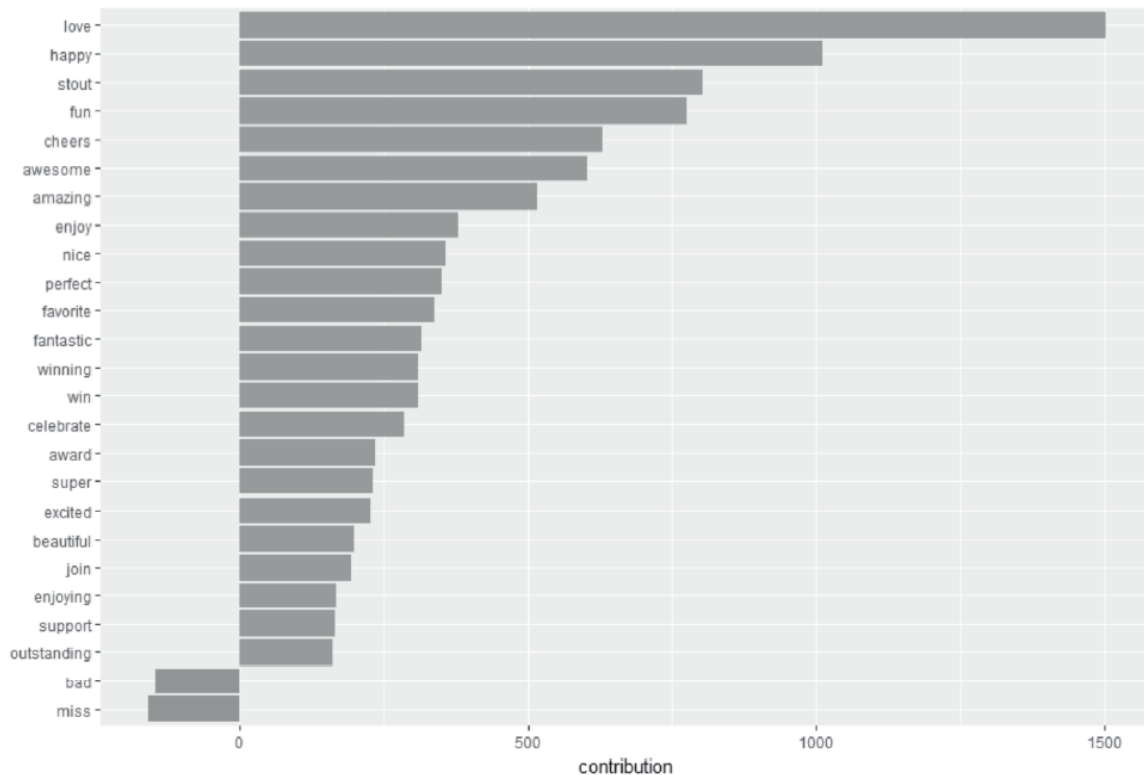


Figure 4: Results of applying the AFINN lexicon to the #craftbeer dataset.

4. Discussion and conclusion

This study described a case study about how text mining methods and sentiment analysis were applied to obtain external knowledge from user-generated content in Twitter to support an entrepreneurial company’s internationalization process. The study proposes that digital channels provide a novel source of market knowledge and can hence provide knowledge about foreign markets, which can be particularly fruitful during the initial stages of firm internationalization.

4.1 Theoretical implications

Fletcher and Harris (2012) studied what new knowledge smaller firms need as they learn to internationalize, and which specific information sources were used by small firms to acquire knowledge during the internationalization process. As smaller firms may not have relevant experience or established networks, other sources of knowledge acquisition become more relevant to them (Fletcher and Harris, 2012). Hence, the present study contributes to the existing body of literature concerning knowledge acquisition for the internationalization of the firm (Brennan and Garvey, 2009; Costa, Soares and Sousa, 2016; Fletcher and Harris, 2012; Mejri and Umemoto, 2010) and suggests that knowledge obtained from social media platforms can provide a new source of indirect experiential knowledge that has not been considered by the previous research.

Previous studies mainly discuss the following types of indirect experiential knowledge: 1) vicarious knowledge, i.e., learning from the experiences of others (Åkerman, 2015; Brennan and Garvey, 2009; Huber, 1991; Schwens and Kabst, 2009) and 2) grafting, referring to knowledge acquisition through acquiring firms and business units or hiring personnel with experiential knowledge (Åkerman, 2015; Huber, 1991; Mejri, Alexander and Tsagdis, 2018). This study introduces social media platforms as novel and potential sources for extracting indirect experiential knowledge, especially for often resource-scarce entrepreneurial firms seeking internationalization. Hult (2003) defines knowledge as credible information that is potentially valuable to an organization. Hence, data obtained from social media and analyzed using data mining methods can become a new source of knowledge and can provide information about the customers, competitive environment and various stakeholders operating in international markets. We propose that this kind of knowledge can be fruitful in dealing with some of the key questions which are relevant to internationalization, such as the selection of foreign markets and the selection of ways of launching into those markets.

In the light of the existing research, the study suggests that knowledge obtained from social media can be particularly fruitful during the initial stages of internationalization. In the model of the influence of knowledge on the internationalization by Casillas, Barbero and Sapienza, (2009), the most typical sources of initial knowledge in terms of internationalization include 1) the individual experiences of the company founder and top management team members, together with 2) the communication between the members of the managerial team, and along with 3) the various memberships these key members have in different organizational networks. It is relatively well-established in the literature (Casillas, Barbero and Sapienza, 2009; Costa, Soares and Sousa, 2016; Hsu, Chen and Cheng, 2013; Muzychenko and Liesch, 2015; Nordman and Melén, 2008; Park, LiPuma and Prange, 2015) that in small and medium-sized enterprises, the knowledge and experience of the firm's founder, i.e., the entrepreneur, indeed plays a key role in the firm's internationalization. However, it is additionally important for a firm to have the ability to acquire knowledge from other sources (Park, LiPuma and Prange, 2015), and to be able to develop by combining and converting knowledge (Fletcher and Harris, 2012). Hence, knowledge obtained both directly and indirectly (Åkerman, 2015) about the international markets effectively complements the prior knowledge held by the entrepreneur and key management team.

4.2 Managerial relevance

This study illustrates a case study of an entrepreneurial firm in which the researchers utilized sentiment analytics methods together with data obtained from social media to increase a firm's knowledge in terms of internationalizing into foreign markets. The findings from this study suggest that in their attempts to internationalize, firms can benefit from the knowledge obtained from social media, and this provides them with a novel source of indirect experiential knowledge which has not been available before. This knowledge can be fruitful for small-and-medium sized firms, because compared to large firms and multinational enterprises, they are less likely to have established organizational routines that need to be unlearned. Further, the study demonstrates a relatively easy-to-implement and cost-effective method for obtaining tacit knowledge from data that is publicly available on social media platforms. Depending on the level of analysis of the data, these methods can provide firms with a better understanding of the market environment in general, or more specific knowledge about customers, competitors, intermediaries and other stakeholders operating in the international supply chain.

For the case company, the study provided generic knowledge about customer attitudes and sentiments regarding craft beers in the pre-specified target markets, and showed that craft beer related Twitter discussions tended to be positively rather than negatively oriented and hence there may exist the potential for the products of the case company. In addition, the results of the study provided knowledge about two competing brands and allowed the firm to compare how customer sentiments differed between the known brands operating in the target markets. With the described methods, firms can extract knowledge about competing products and brands in international markets, compare the image of competing firms in terms of the sentiments that customers associate with them, and extract knowledge of how competitors, intermediaries and other stakeholders engage on social networks by applying social media network analytic methods.

Previous research highlights that experiential knowledge, in general, tends to be situation-specific and costly to obtain directly (Mejri , Alexander and Tsagdis, 2018), and hence this study proposes that indirect experiential knowledge obtained from social media could be a feasible option particularly during the initial phases of internationalization. Åkerman (2015) examined different knowledge-acquisition strategies among internationalizing firms, and based on the results, firms with a diversified knowledge-acquisition strategy held more market information than the other groups compared and therefore had a better overall understanding of the international markets. Further, Armario, Ruiz and Armario (2008) showed that firms acknowledging the relevance of information about customers and competitors when designing their international strategies tend to achieve a better overall performance in foreign markets. The managerial implications of this study, therefore, suggest that firms should use diversified knowledge-acquisition strategies in their attempts to enter international markets, and innovative methods not considered before can provide new and fruitful avenues to support the other methods companies use to obtain knowledge about international markets.

4.3 Limitations and future research

The study illustrated a case study in which data obtained from Twitter was used to extract market knowledge to support company plans to internationalize. Rather than providing widely generalizable results, the main contribution of the study is in demonstrating how companies could apply novel content analysis methods and knowledge extracted from publicly available Twitter data.

To keep the research design simple, we illustrated a case in which content analysis methods were used to obtain indirect experiential market knowledge during the initial phases of internationalization of a firm. The methods were based on word frequencies and sentiments derived from the content (Chae, 2015), and such methods should be reinforced with the use of other data analysis methods such as descriptive analytics (e.g., tweet and user metrics) and network analytics (e.g., community, centrality and topology analyses). Future research could extend our findings and study the advantages and limitations of content analytics, descriptive analytics and network analytics during different phases of a firm's internationalization process, including pre-internationalization, the start of internationalization, novice internationalization, experienced internationalization etcetera. Network knowledge (i.e., knowledge about social and business networks that facilitate the internationalization of the firm) is proposed to be the most useful type of knowledge from the pre-internationalization phase until the novice internationalization phase (Mejri and Umemoto, 2010), and future studies could look into how social media network analytics methods, for instance, contribute to knowledge acquired by the firm's direct experience in international markets.

The study suggests that indirect experiential knowledge obtained from social media can be a novel and beneficial source of market knowledge – however – we consider it as an additional and complementary source of knowledge in addition to other types of knowledge existing in the literature (Åkerman, 2015; Huber, 1991; Mejri, Alexander and Tsagdis, 2018; Nguyen et al., 2015). Hence, future research should broaden the scope of the data in order to provide more generalizable findings. With regard to future research, the present study focuses solely on the data obtained from Twitter. However, data obtained from other social media platforms could have been utilized as well, which provides a potential avenue for future research. Additionally, in this study we focused on the use of textual data together with selected sentiment analytics methods. Future studies are encouraged to pursue other methods of data mining to extract knowledge from other forms of data, including image, video and voice data.

References

- Åkerman, N., 2015. Knowledge-acquisition strategies and the effects on market knowledge - profiling the internationalizing firm. *European Management Journal*, 33(2), pp. 79–88.
- Armario, J. M., Ruiz, D. M., and Armario, E. M., 2008. Market orientation and internationalization in small and medium-sized enterprises. *Journal of Small Business Management*, 46(4), pp. 485–511.
- Balahur, A., and Jacquet, G. 2015. Sentiment analysis meets social media - challenges and solutions of the field in view of the current information sharing context. *Information Processing and Management*, 51(4), pp. 428–432.
- Balducci, B., and Marinova, D. 2018. Unstructured data in marketing. *Journal of the Academy of Marketing Science*, 46(4), pp. 557–590.
- Barnes, N. G., Kane, A., and Maloney, K. 2018. The 2018 Fortune 500 target millennials and seek uncensored expression. <https://www.umassd.edu/cmr/social-media-research/2018-fortune-500/#d.en.1535637> [Accessed 20.2.2019].
- Berthon, P. R., Pitt, L. F., Plangger, K., and Shapiro, D. 2012. Marketing meets web 2.0, social media, and creative consumers: implications for international marketing strategy. *Business Horizons*, 55(3), pp. 261–271.
- Bose, R. 2008. Competitive intelligence process and tools for intelligence analysis. *Industrial Management & Data Systems*, 108(4), pp. 510–528.
- Bougie, R., Pieters, R., and Zeelenberg, M. 2003. Angry customers don't come back, they get back: the experience and behavioral implications of anger and dissatisfaction in services. *Journal of the Academy of Marketing Science*, 31(4), pp. 377–393.
- Bradley, M. M., and Lang, P. J. 2008. Affective norms for english words (ANEW): instruction manual and affective ratings. *IEEE Internet Computing*, 12(5), pp. 44–52.
- Bravo-Marquez, F., Mendoza, M., and Poblete, B. 2014. Meta-level sentiment models for big social data analysis. *Knowledge-Based Systems*, 69(1), pp. 86–99.
- Brennan, L., and Garvey, D. 2009. The role of knowledge in internationalization. *Research in International Business and Finance*, 23(2), pp. 120–133.
- Buckley, P. J., Munjal, S., Enderwick, P., and Forsans, N. 2016. The role of experiential and non-experiential knowledge in cross-border acquisitions: the case of Indian multinational enterprises. *Journal of World Business*, 51(5), pp. 675–685.
- Cao, Q., Duan, W., and Gan, Q. 2011. Exploring determinants of voting for the "helpfulness" of online user reviews: a text mining approach. *Decision Support Systems*, 50(2), pp. 511–521.
- Carr, C. T., and Hayes, R. A. 2015. Social media: defining, developing, and divining. *Atlantic Journal of Communication*, 23(1), pp. 46–65.
- Casillas, J. C., Barbero, J. L., and Sapienza, H. J. 2015. Knowledge acquisition, learning, and the initial pace of internationalization. *International Business Review*, 24(1), pp. 102–114.
- Casillas, J. C., Moreno, A. M., Acedo, F. J., Gallego, M. A., and Ramos, E. 2009. An integrative model of the role of knowledge in the internationalization process. *Journal of World Business*, 44(3), pp. 311–322.

- Cesaroni, F. M., and Consoli, D. 2015. Are small businesses really able to take advantage of social media? *Electronic Journal of Knowledge Management*, 13(4), pp. 257-268.
- Chae, B. 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, pp. 247-259.
- Costa, E., Soares, A. L., and Sousa, J. P. De. 2016. Information, knowledge and collaboration management in the internationalisation of smes. *International Journal of Information Management*, 36(4), pp. 557-569.
- De Clercq, D., Sapienza, H. J., Yavuz, R. I., and Zhou, L. 2012. Learning and knowledge in early internationalization research: past accomplishments and future directions. *Journal of Business Venturing*, 27(1), pp. 143-165.
- De Vries, L., Gensler, S., and Leeflang, P. S. H. 2012. Popularity of brand posts on brand fan pages: an investigation of the effects of social media marketing. *Journal of Interactive Marketing*, 26(2), pp. 83-91.
- DeCarolis, D. M., and Deeds, D. L. 1999. The impact of stocks and flows of organizational knowledge on firm performance: an empirical investigation of the biotechnology industry. *Strategic Management Journal*, 20(10), pp. 953-968.
- Deligianni, I., Voudouris, I., and Lioukas, S. 2015. Growth paths of small technology firms: the effects of different knowledge types over time. *Journal of World Business*, 50(3), pp. 491-504.
- Dey, L., Haque, S. M., Khurdiya, A., and Shroff, G. 2011. Acquiring competitive intelligence from social media. In *Proceedings of the 2011 Joint Workshop on Multilingual OCR and Analytics for Noisy Unstructured Text Data*.
- Eriksson, K., Johanson, J., Majkgård, A., and Sharma, D. D. 1997. Experiential knowledge and costs in the internationalization process. *Journal of International Business Studies*, 28(2), pp. 337-360.
- Eriksson, K., Johanson, J., Majkgård, A., and Sharma, D. D. 2000. Effect of variation on knowledge accumulation in the internationalization process. *International Studies of Management & Organization*, 30(1), pp. 26-44.
- Fleisher, C. S. 2008. Using open source data in developing competitive and marketing intelligence. *European Journal of Marketing*, 42(7/8), pp. 852-866.
- Fletcher, M., and Harris, S. 2012. Knowledge acquisition for the internationalization of the smaller firm: content and sources. *International Business Review*, 21(4), pp. 631-647.
- Fraccastoro, S., and Gabrielsson, M. 2018. Effects of social media usage on the speed of international growth of an international new venture. In *Proceedings of the 51st Hawaii International Conference on System Sciences*, pp. 4309-4316.
- Freeman, S., Hutchings, K., Lazaris, M., and Zyngier, S. 2010. A model of rapid knowledge development: the smaller born-global firm. *International Business Review*, 19(1), pp. 70-84.
- García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M., and Herrera, F. 2016. Big data preprocessing: methods and prospects. *Big Data Analytics*, 1(1), pp. 1-22.
- Gassmann, O., and Keupp, M. M. 2007. The competitive advantage of early and rapidly internationalising smes in the biotechnology industry: a knowledge-based view. *Journal of World Business*, 42(3), pp. 350-366.
- Gaur, S. S., Herjanto, H., and Makkar, M. 2014. Review of emotions research in marketing, 2002-2013. *Journal of Retailing and Consumer Services*, 21(6), pp. 917-923.
- Grant, R. M. 1996. Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17, pp. 109-122.
- Han, J., and Kamber, M. 2001. *Data mining: Concepts and techniques*. San Francisco: Morgan Kaufmann.
- Haverila, M., and Ashill, N. 2011. Market intelligence and NPD success: a study of technology intensive companies in Finland. *Marketing Intelligence & Planning*, 29(5), pp. 556-576.
- He, W., Zha, S., and Li, L. 2013. Social media competitive analysis and text mining: a case study in the pizza industry. *International Journal of Information Management*, 33(3), pp. 464-472.
- Hilmersson, M. 2014. Experiential knowledge types and profiles of internationalising small and medium-sized enterprises. *International Small Business Journal: Researching Entrepreneurship*, 32(7), pp. 802-817.
- Hsu, W. T., Chen, H. L., and Cheng, C. Y. 2013. Internationalization and firm performance of smes: the moderating effects of ceo attributes. *Journal of World Business*, 48(1), 1-12.
- Huber, G. P. 1991. Organizational learning: the contributing processes and the literatures. *Organization Science*, 2(1), pp. 88-115.
- Hult, G. T. M. 2003. An integration of thoughts on knowledge management. *Decision Sciences*, 34(2), pp. 189-195.
- Johanson, J., and Vahlne, J. E. 1977. Process of the firm—a model of knowledge foreign and increasing market commitments. *Journal of International Business Studies*, 8, pp. 23-32.
- Katsikea, E., Theodosiou, M., and Makri, K. 2019. The interplay between market intelligence activities and sales strategy as drivers of performance in foreign markets. *European Journal of Marketing*, In press.
- Kim, J., and Hastak, M. 2018. Social network analysis: characteristics of online social networks after a disaster. *International Journal of Information Management*, 38(1), pp. 86-96.
- Kohli, A. K., and Jaworski, B. J. 1990. Market orientation: the construct, research propositions, and managerial implications. *Journal of Marketing*, 54(2), pp. 1-18.
- Kuivalainen, O., Saarenketo, S., and Puumalainen, K. 2012. Start-up patterns of internationalization: a framework and its application in the context of knowledge-intensive smes. *European Management Journal*, 30(4), pp. 372-385.
- Lord, M. D., and Ranft, A. L. 2000. Organizational learning international internal markets: about exploring new the market transfer of local knowledge. *Journal of International Business Studies*, 31(4), pp. 573-589.
- Maltby, T. 2012. Using social media to accelerate the internationalization of startups from inception. *Technology Innovation Management Review*, 2(10), pp. 22-26.

- Maltz, E., and Kohli, A. K. 1996. Market intelligence dissemination across functional boundaries. *Journal of Marketing Research*, 33(1), pp. 47-61.
- Mejri, K., Alexander, J., and Tsagdis, D. 2018. Knowledge configurations of small and medium-sized knowledge-intensive firms in a developing economy: a knowledge-based view of business-to-business internationalization. *Industrial Marketing Management*, 71, pp. 160-170.
- Mejri, K., and Umemoto, K. 2010. Small- and medium-sized enterprise internationalization: towards the knowledge-based model. *Journal of International Entrepreneurship*, 8(2), pp. 156-167.
- Mohammad, S. M., and Turney, P. D. 2013. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3), pp. 436-465.
- Mostafa, M. M. 2013. More than words: social networks' text mining for consumer brand sentiments. *Expert Systems with Applications*, 40(10), pp. 4241-4251.
- Muzychenko, O., and Liesch, P. W. 2015. International opportunity identification in the internationalisation of the firm. *Journal of World Business*, 50(4), pp. 704-717.
- Nguyen, B., Yu, X., Melewar, T. C., and Chen, J. 2015. Brand innovation and social media: knowledge acquisition from social media, market orientation, and the moderating role of social media strategic capability. *Industrial Marketing Management*, 51, pp. 11-25.
- Nielsen, F. Å. 2011. A new ANEW: evaluation of a word list for sentiment analysis in microblogs. *CEUR Workshop Proceedings*, 718, pp. 93-98.
- Nonaka, I. 1994. A dynamic theory of organizational knowledge creation. *Organization Science*, 5(1), pp. 14-37.
- Nordman, E. R., and Melén, S. 2008. The impact of different kinds of knowledge for the internationalization process of born globals in the biotech business. *Journal of World Business*, 43(2), pp. 171-185.
- Okazaki, S., and Taylor, C. R. 2013. Social media and international advertising: theoretical challenges and future directions. *International Marketing Review*, 30(1), pp. 56-71.
- Oliveira, N., Cortez, P., and Areal, N. 2016. Stock market sentiment lexicon acquisition using microblogging data and statistical measures. *Decision Support Systems*, 85, pp. 62-73.
- Oviatt, B. M., and McDougall, P. P. 2005. Defining international entrepreneurship and modeling the speed of internationalization. *Entrepreneurship Theory and Practice*, 29(5), pp. 537-553.
- Park, S., LiPuma, J. A., and Prange, C. 2015. Venture capitalist and entrepreneur knowledge of new venture internationalization: a review of knowledge components. *International Small Business Journal: Researching Entrepreneurship*, 33(8), pp. 901-928.
- Pathak, X., and Pathak-Shelat, M. 2017. Sentiment analysis of virtual brand communities for effective tribal marketing. *Journal of Research in Interactive Marketing*, 11(1), pp. 16-38.
- Pellegrino, J. M., and McNaughton, R. B. 2017. Beyond learning by experience: the use of alternative learning processes by incrementally and rapidly internationalizing SMEs. *International Business Review*, 26(4), pp. 614-627.
- Plutchik, R. 2002. *Emotions and life: perspectives from psychology, biology and evolution*. Washington, DC: American Psychological Association.
- Rello, L., & Baeza-Yates, R. 2012. Social media is not that bad! The lexical quality of social media. In *Sixth International AAAI Conference on Weblogs and Social Media*.
- Schoenherr, T., and Swink, M. 2015. The roles of supply chain intelligence and adaptability in new product launch success. *Decision Sciences*, 46(5), pp. 901-936.
- Schwens, C., and Kabst, R. 2009. How early opposed to late internationalizers learn: experience of others and paradigms of interpretation. *International Business Review*, 18(5), pp. 509-522.
- Sharma, P., and Chrisman, S. J. J. 2007. Toward a reconciliation of the definitional issues in the field of corporate entrepreneurship. *Entrepreneurship*, 83, pp. 83-103.
- Singh, T., and Kumari, M. 2016. Role of text pre-processing in twitter sentiment analysis. *Procedia Computer Science*, 89, pp. 549-554.
- Stoian, M. C., Dimitratos, P., and Plakoyiannaki, E. 2018. SME internationalization beyond exporting: a knowledge-based perspective across managers and advisers. *Journal of World Business*, 53(5), pp. 768-779.
- Tang, G., Pei, J., and Luk, W. S. 2014. Email mining: tasks, common techniques, and tools. *Knowledge and Information Systems*, 41(1), pp. 1-31.
- Tsimonis, G., and Dimitriadis, S. 2014. Brand strategies in social media. *Marketing Intelligence and Planning*, 32(3), pp. 328-344.
- Vasileiadou, E. M., and Missler-Behr, M. 2011. Virtual embeddedness and social media as a basis for the relational capital management of new ventures. *Electronic Journal of Knowledge Management*, 9(3), pp. 188-203.
- Wernerfelt, B. 1984. A resource-based view of the firm. *Strategic Management Journal*, 5(2), pp. 171-180.
- Wickham, H. 2014. Tidy data. *Journal of Statistical Software*, 59(10).
- Yeniyurt, S., Cavusgil, S. T., and Hult, G. T. M. 2005. A global market advantage framework: the role of global market knowledge competencies. *International Business Review*, 14(1), pp. 1-19.
- Zhou, L. 2007. The effects of entrepreneurial proclivity and foreign market knowledge on early internationalization. *Journal of World Business*, 42(3), pp. 281-293.