Article

Developing a Digital Artifact for the Sustainable Presentation of Marketing Research Results

Zheng Shen ^{1,2,*} and Armida de la Garza ²¹ Department of Management & Marketing, University College Cork, Cork T12 K8AF, Ireland² Department of Digital Arts and Humanities, University College Cork, Cork T12 K8AF, Ireland; adelagarza@ucc.ie

* Correspondence: 115220146@umail.ucc.ie

Received: 18 October 2019; Accepted: 18 November 2019; Published: 20 November 2019

Abstract: The rapid development of technology transforms the way researchers conduct projects, communicate with others, and disseminate findings. In addition to traditional presentations of research results, this paper argues that building a digital artifact is another optional method for the dissemination of research findings from the perspective of marketing. Thus, 20 Irish and Chinese micro-influencers were investigated from March 2016 to March 2019, and their microblogs were analyzed by text mining techniques. Consequently, the paper finds four types of keywords that micro-influencers apply to their marketing on social media. Based on the marketing keywords, a digital tool is designed to label fashion keywords in the microblogging automatically. The proposed tool not only contributes to model fashion bloggers' content and increase the influence of marketing on social media but also enlightens marketing scholars to develop digital tools for the sustainability of disseminating research results.

Keywords: digital artifact; sustainability; marketing; social media; fashion microblogging; text mining

1. Introduction

With the invention of Web 2.0 in 1999, social media starts to play an influential role in our daily life nowadays. It changes the form we “interact, play, shop, read, write, work, listen, create, communicate, collaborate, produce, co-produce, search, and browse” [1]. For example, consumers prefer to buy fashion products online rather than in-store. In China, the profit of online shopping is estimated at over 150 billion CNY [2]. Likewise, the revenue in Ireland 2019 amounts to €680m with an annual growth rate of 17.3% [3]. At present, consumers admit that they make final purchase decisions with the assistance of social media [4]. Specifically, they gather fashion information, read others' blogs, communicate with other consumers, and express their experiences of purchase. The surveys of Angella J. Kim et al. and Amanda Lenhart et al. indicate that 63% of adults get information from social media, more than 73% of adults use Facebook to blog their status such as what they are wearing, and 60% of online fashion consumers incline to communicate their experiences on brands and products with others through social media [4,5]. Consequently, “social media content has been used by various brands for competing with the competitors, promoting products and offers, and maintaining a reputation” [6]. As a number of brands engaged with social media content, Sonja Jefferson and Sharon Tanton argue that the key to the success of social media marketing is how to make quality content [7]. The high-quality content can help outstanding brands in amounts of content marketing, and enhance the marketing influence. For this reason, the enterprise marketing departments are advised to use smart content to impact consumers on social media [8].

In content marketing, influencer-generated content can further increase marketing significance. Referring to Robert V. Kozinets et al., consumers can be influenced by members of the consumer

network through exchanging marketing messages deliberately and directly [9]. When the influence of messages is overwhelming, the members of the consumer network should be noticed because they turn out to be influencers. Theo Araujo et al. analyzed over 5300 tweets to figure out the role of influential individuals for branding [10]. The result shows that the influence of brand messages hugely depends on the number of influencers who retweet the messages. Hence, nowadays companies realize the significance of influencers, try their best to find power-users or people who already have a significant effect on the social network, and collaborate with them for targeting potential consumers [11]. However, Christian Hughes et al. argue that “influencer marketing is prevalent in firm strategies, yet little is known about the factors that drive the success of online brand engagement at different stages of the consumer purchase funnel” [12]. As a result, the content analysis in the research reveals the successful factors of influencers, and make up for previous studies.

So as to examine micro-influencers, currently there are three main approaches: User Attributes Analysis, Network Structure Analysis and Text Mining Analysis. The User Attributes Analysis concentrates on influencers’ individual characteristics. For instance, Gabriela Ramirez-de-la-Rosa et al. suggest examining users’ writing styles and behaviors to identify opinion leaders on Twitter [13]. In terms of Network Structure Analysis, studies can be further categorized into two trends. The first trend is to discover influencers based on the classical network typology analysis, and another trend is to look for leaders by means of Social Network Analysis [14,15]. The frequently used methods consist of questionnaires and content analysis. Last but not least, Text Mining Analysis aims to find influencers on the large scale of social media networks through automated computational techniques. By comparison, Text Mining Analysis is more appropriate for this study. Due to the popularity of social media, innumerable data are produced every day. The large quantity of data causes the difficulties of analysis like User Attributes Analysis and Network Structure Analysis. However, the automated text mining techniques are considered to help researchers identify features of leaders in the large scale of social networks efficiently [16]. More importantly, it can not only recognize influencers but also deal with social media content. Sofus A. Macskassy’s research certifies that text mining enables the topic-based analysis of blogging and finding influencers in large social networks [17]. Thus, the study applies text mining to analyze influencers and their marketing content in the microblogging.

Also, social media alters scholars’ communicative patterns of research in the academy. Previously, book and journal publishing took prominent roles for researchers to communicate research results with others [18]. At present, it is argued that publication patterns in the social sciences and humanities should be diverse [19]. One of the most successful stories is Jack A. Heinemann et al.’s research paper on agricultural sustainability. It was published on Twitter, retweeted by 496 accounts and viewed over 8000 times in two weeks. The fast dissemination indicates that a social media platform like Twitter is an ideal venue for researchers to engage with public audiences and transcend the traditional venues for academic knowledge dissemination [20,21]. As a result, recent research has attempted to study the prevalence, volume, and meaning of sharing of research on various platforms [22]. In such a case, this paper further urges marketing researchers to create digital artifacts as an alternative method of communicating research results. Correspondingly, scholars begin to admit that an approach based on artificial intelligence already is vital to marketing and is used increasingly [23]. Furthermore, artificial intelligence is “paving the way for the future of marketing and business transformation” [24]. Therefore, this paper addresses the analysis of content marketing in micro-influencers’ fashion microblogs by means of text mining at first. Afterward, it presents how the results are applied to develop the digital artifact for disseminating findings sustainably.

2. Smart Content Marketing

Content marketing has been a top strategy for many years. Successful content marketing can increase brand visibility, drive traffic to websites, help educate and convert customers [25]. With the popularity of social media, blogging becomes a significant channel for content marketing. However, content marketing in blogging has a problem of oversaturation at present. Marketers’ blogs are

considered to be “the templated, mass-distributed messaging of the past” [26]. In other words, marketers are too lazy to check others’ viewpoints in blogging and copy articles simply [27]. As a consequence, a large number of blogs on social media turn out to be monotonous and meaningless from the perspective of marketing, because they hardly convince consumers. Hence, current marketers are eager to distinguish themselves from others on social media utilizing smart content marketing. The smart content marketing is consumer-oriented, innovative and interactive. The highly targeting and segmenting content for audiences remains one of 10 powerful marketing tools and tactics that shake up the industry in 2019 [28]. For this reason, the study focuses on the examination of Irish and Chinese micro-influencers’ content marketing, and reveals the successful factors of their influence on marketing. well-established methods can be briefly described and appropriately cited.

In terms of content marketing, a keyword is the core of smart content marketing. Marketers can take advantage of critical terms to help marketing content appear online frequently. Rebecca Lieb claims that keywords are crucial for content marketing and Search Engine Optimization (SEO) [29]. That is to say, consumers search for information and receive relevant information on the ground of keywords. For instance, Kinshuk Jerath et al. investigated the relations between keyword popularity and consumers’ click behaviors [30]. The result shows that keywords affect consumers’ receiving content online and lead them to click on sponsored links. As a result, keywords benefit from reaching target consumers at the right time. Marketers can motivate consumers through critical terms. At the same time, consumers are not missed in the social network if marketers optimize their content based on keywords. Andrey Simonov and Chris Nosko analyzed how focal brands use keywords to compete with other relevant firms [31]. The research result finds that competitors can steal 10–20% of clicks on average when focal brands are not shown in the top rank of keyword searching. Otherwise, competitors can steal merely 1–5% of clicks when focal brands are top ranks. Thus, keywords contribute to the traffic of content marketing and superiority in the competitions.

More importantly, traffic influence can affect the final sales. Shijie Lu and Sha Yang conducted a study on the influence of keyword market entries in sponsored search advertising. The result indicates that “the keyword-specific competition information provided by infomediaries can improve the search engine’s revenue by about 5.7%” [32]. In other words, keywords assist marketers in defeating their competitors by means of top ranking on the search engines and increasing online marketing revenues. In particular, keywords in content marketing are essential in the current era of big data. Among tons of posts every day, how to make a specific blog stand out for drawing consumers’ attention is a serious question for digital marketers. In order to answer this question, the proper keyword selection in developing the content of blog marketing tends to be the right solution. Supported by Arokia R. Terrance et al., the website developer should apply keyword analysis to digital marketing in order to rank the content result in the first place of search engines [33]. In short, the top rank of content enables the high visibility for consumers online. Eventually, it increases the traffic of consumers and the overall sales of products. As a result, this study concentrates on the identification of keywords for the development of a digital artifact on micro-influencers’ content marketing.

Referring to the Content Marketing Institute, 62% of the most influential content marketers have a documented strategy [34]. Varieties of influencers’ strategies make other marketers hardly to perceive the pattern of content marketing in the short term. The analyzing increased diversity and volume of content marketing strategies are beyond the competence of the human mind [35]. Thus, it not only urges to develop smart content for social media marketing but also finds an appropriate way to detect the model of content marketing. Consequently, this study applies the computer-assisted method—text mining analysis to help understand micro-influencers’ smart content marketing on social media.

3. Text Mining Methods

Nowadays, it is estimated that people generate 2.5 Exabytes data (1 Exabyte = 1,000,000 Terabytes) every day [36]. This incredible growth of data is considered mainly from social media posts [37]. According to Wenbo Wang et al., Twitter produces at an enormous speed of 340 million posts every

day [38]. As a result, big data challenges marketers to understand, use, store, and present. Ramzan Talib et al. compare a variety of techniques, and point out that “the selection of right and appropriate text mining technique helps to enhance the speed and decreases the time and effort required to extract valuable information” [39]. Text mining is considered as one of the best practices because it can find predictive patterns for both structured and unstructured texts [40]. It offers an alternative method to collect market insights [41]. Hence, text mining benefits to derive high-quality information from a large scale of data and discover the pattern of content marketing on social media efficiently.

For years, text mining has been conducted in a broad range of fields like healthcare [42,43], politics [44,45], arts [46,47], and education [48,49]. Concerning social media marketing, Mohamed M. Mostafa investigated 3516 tweets for analyzing consumers’ sentiments on global brands by text mining techniques, and reveal the value of using text mining in studies on blogging and social media [50]. Besides, Aron Culotta and Jennifer Culter used their research to mine brand perceptions from 200 brands ranging from apparel and cars to food and personal care [51]. In comparison to costly as well as time-consuming traditional methods, the research proves that text mining is certified to be a novel, general, automated, reliable, flexible, and scalable approach to monitor brand perceptions, and understand brand-consumer relationships on social media. As a consequence, this study employs text mining to explore content marketing in fashion microblogging at first, and then develop a digital artifact to present research results.

3.1. Data Collection

The data come from fashion microblogs written by 20 Irish and Chinese bloggers from March 2016 to March 2019. The number of fashion bloggers is enormous and growing every day. Thus, not all fashion bloggers in Ireland and China can be studied at one time. For this reason, the study concentrates on fashion micro-influencers who have the most influence on consumers through social media marketing. For measuring the influence of social media activities on consumers, Jeremiah Owyang from Altimeter Group and John Lovett from Web Analytics Demystified suggest utilizing Key Performance Indicators (KPIs) [52]. They conclude four measurement frameworks—Foster Dialog, Promote Advocacy, Facilitate Support and Spur Innovation—in line with business objectives. Among these four measurement frameworks, Promote Advocacy is the framework closely related to the measurement of influence, which “allows businesses to extend their reach beyond their immediate circles of influence by taking advantage of word of mouth and viral activity” [11]. It has three Key Performance Indicators—Active Advocates, Advocate Influence as well as Advocacy Impact. Among three KPIs, Advocate Influence is chosen for the project because it can indicate “the unique advocate’s influence across one or more social media channels” [11]. The influence is measured by the number of comments, reach, relevant contents and shares. The active influence is calculated by dividing a single advocate’s influence by the total number of advocates (see the following equation):

$$\text{Active Influence} = \text{Unique Advocate's Influence} / \text{Total Advocate's Influence}, \quad (1)$$

In order to calculate the active influence of micro-influencers, we investigated the lists of most influential bloggers in Ireland and China for determining the ranges in the selection at the beginning. According to their volume of comments, reach, relevant contents and shares of fashion microblogs, consequently 20 most influential Irish and Chinese micro-influencers were chosen for this study. The results are shown in Tables 1 and 2.

Table 1. Top 10 Irish Micro-influencers.

Irish Micro-Influencers	No. of Comments	No. of Reach	No. of Relevant Contents	No. Shares	Active Influence
Sosueme	4016	605,207	11,974	5663	0.88
Thunder and Threads	3512	569,453	11,855	4540	0.82
Pippa	3244	493,959	11,547	4505	0.81
Help my style	3303	424,584	9255	4132	0.78
Anouska	3453	389,824	8090	3949	0.7
Fluff and Fripperies	3371	390,842	7998	3988	0.7
The Style Fairy	3200	276,472	6367	3323	0.63
What she wears	3169	276,228	6163	3585	0.63
Just Jordan	2179	195,426	5222	2340	0.57
Love Lauren	2112	174,200	5230	2255	0.53

Table 2. Top 10 Chinese Micro-influencers.

Chinese micro-Influencers	No. of Comments	No. of Reach	No. of Relevant Contents	No. Shares	Active Influence
Shiliupobaogao	22,630	9,534,440	41,340	81,980	0.94
Yang Fan Jame	36,500	9,077,270	31,908	61,790	0.89
Han Huohuo	14,892	9,028,789	31,056	51,914	0.87
Chrison	25,404	8,884,001	30,156	62,791	0.86
Peter Xu	23,215	8,490,286	23,372	51,134	0.77
Gogoboi	24,528	8,320,808	18,945	29,200	0.72
Mr. Kira	12,410	5,337,882	16,733	12,118	0.7
Qiangkouxiaolajiao	29,200	5,227,843	13,663	25,477	0.69
Miss Shopping Li	15,630	3,605,081	10,079	20,300	0.68
Boy Mr. K	17,305	1,342,843	11,631	24,090	0.51

Influencers are defined as a “third party who significantly shapes the customer’s purchasing decision” [53]. In business marketing, representative influencers include industry analysts, consultants, and journalists. With the development of technologies, influencers are not limited to these occupations. For instance, Thunder and Threads is a college student, and Help my style is a TV presenter. They are keen on using social media, especially microblogging, to communicate with other members of the network and achieve a significant influence on them. Also, all of them have a large number of followers on social media compared with other bloggers. Sophie C. Boerman defines micro-influencers as “‘normal’ people who turned Instafamous and typically have dozens to hundreds of followers” [54]. Tables 1 and 2 show that they have a considerable number of reach in the social network. Therefore, they can be further identified as micro-influencers, who significantly shape the purchasing decision of consumers in the same social network through social media marketing.

In relation to online fashion marketing, micro-influencers are featured by loving fashion and specializing in fashion. Referring to Sosueme, she has started to microblog since 2010 because she is very interested in fashion. The other Irish micro-influencers also began microblogging in 2009 and 2013. By comparison, Chinese micro-influencers have started earlier. Most of them started in 2007 because of their jobs. For example, Han Huohuo and Gogoboi work as fashion editors. One part of their work is to read fashion news abroad and introduce it to Chinese consumers. Hence, fashion microblogging becomes a channel for them to diffuse fashion and influence consumers’ purchase behaviors. Besides, these 20 micro-influencers indicate new characteristics of fashion micro-influencers. For one thing, Chinese fashion micro-influencers are more masculine than Irish micro-influencers. In the study, only one Irish micro-influencer is a man while eight out of ten Chinese micro-influencers are men. The result shows the difference from previous studies that prove fashion influencers are mostly females [55,56]. For another, most of 20 influencers are in between the thirties and forties, which are not young described in the previous research [57,58]. As a result, fashion micro-influencers can be described as middle-aged, loving fashion, expertizing in fashion, owning a large number of followers

and reach in the social network, diffusing fashion on social media, and influencing other members of the same network significantly.

3.2. Data Analysis

In general, the techniques of text mining consist of information retrieval, natural language processing, information extraction from text, text summarization, unsupervised learning methods, supervised learning methods, probabilistic methods, text streams, and social media mining, opinion mining and sentiment analysis, and biomedical text mining [59]. Among them, Natural Language Processing (NLP) is the core of text mining. It contains word segmentation, part-of-speech (POS) tagging, named entity recognizers (NER), topic detection and classification, etc. Referring to Atefeh Farzindar and Diana Inkpen, “Natural language processing is one of the most promising avenues for social media data processing” [60]. They indicate three stages for natural language processing—Linguistic Pre-processing, Semantic Analysis, and finally applications.

3.2.1. Linguistic Pre-Processing

The Linguistic Pre-processing includes text normalization, part-of-speech taggers and named entity recognizers for re-training data for the final application. In order to achieve this stage, Natural Language Toolkit (NLTK) is utilized, which is one of the most well-known tools for NLP. “NLTK was originally created in 2001 as part of a computational linguistics course in the Department of Computer and Information Science at the University of Pennsylvania” [61]. According to Bird et al., NLTK is characterized by simplicity, consistency, extensibility, and modularity. More importantly, it supports many NLP techniques such as tokenization, POS tagging, NER, and semantic interpretation. Therefore, the micro-influencers’ microblogs were analyzed by NLTK in the study. Take one of Pippa’s microblogs as an example, and the analysis of NLTK is illustrated step by step as below:

Example:

Pippa O’Connor 30 December 2016

Yesss 🙌😊 Our winter sales on pocobypippa.com and pippacollection.com are ending tomorrow at midnight ✨👁️ They are also happening in the pop-up shop in Dundrum Town Centre too 😊 (From Pippa O’Connor on 30 December 2016)

As seen in the example, the raw content of fashion microblogging contains texts, emojis, dates, bloggers’ names, and errors sometimes. For instance, the word “Yesss” is a modified form of the word “Yes”. For NLTK, the wrong words can be extracted from the content because of spaces between words, but they can not be identified in the POS and NER subsequently. Hence, it is necessary to normalize the microblogging content before text mining analysis. Similarly, it is essential to clear emojis from the data, which cannot be analyzed by NLTK at present. As a result, the example is normalized as follows:

Text Normalization: Yes Our winter sales on pocobypippa.com and pippacollection.com are ending tomorrow at midnight They are also happening in the pop-up shop in Dundrum Town Centre too

After the text normalization, each word is tokenized from fashion microblogs by tokenization of NLTK. The tokenization separates each word in the text data. The correctness of tokenization is fundamental for text mining. Referring to Sonali Vijay Gaikwad et al., “information extraction is the initial step for computer to analyze unstructured text by identifying key phrases and relationships within text” [62]. Therefore, the example is tokenized as below:

NLTK Tokenization:

```
>>> nltk.word_tokenize(text)
```

```
>>> ['Yes', 'Our', 'winter', 'sales', 'on', 'pocobypippa.com', 'and', 'pippacollection.com', 'are',
'ending', 'tomorrow', 'at', 'midnight', 'They', 'are', 'also', 'happening', 'in', 'the', 'pop', 'up',
'shop', 'in', 'Dundrum', 'Town', 'Centre', 'too']
```

In the process of tokenization, it is found that NLTK is a good indicator of tokenizers, especially in English, but it has difficulties in dealing with text data in relation to fashion. For instance, the words “Dundrum”, “Town” and “Centre” are regarded as three tokenizers in the example. From the perspective of semantic analysis, “Dundrum”, “Town” and “Centre” can be considered as one tokenizer. Besides, it is challenging for NLTK to handle tokenization in Chinese because there is no space between Chinese words. In general, researchers try to teach computers to understand Chinese text data based on the comparison with Chinese dictionaries and a large number of previous statistics. That is to say, it is crucial to establish a database of fashion microblog marketing for accurate tokenization in English and Chinese.

Additionally, Part-of-speech (POS) identifies the parts of words taken in the sentences. More concretely, the parts include nouns, verbs, adjectives, adverbs, and conjunctions. Each word of the text data is tagged as these parts respectively. The text data can be tagged by the program code “`nltk.pos_tag (nltk.word_tokenize (text))`” in the NLTK (see the following instance). According to Farzindar and Inkpen, “POS taggers clearly need re-training in order to be usable on social media data. Even the set of POS tags used must be extended in order to adapt to the needs of this kind of text” [60]. Therefore, POS taggers are re-trained for fashion-related content marketing in the study when designing the digital artifact, which is elaborated in the subsequent section.

NLTK POS:

```
>>> nltk.pos_tag (nltk.word_tokenize (text))
>>> [('Yes', 'VB'), ('Our', 'PRP$'), ('winter', 'NN'), ('sales', 'NNS'), ('on', 'IN'),
('pocobypippa.com', 'NN'), ('and', 'CC'), ('pippacollection.com', 'NN'), ('are', 'VBP'), ('ending',
'VBG'), ('tomorrow', 'NN'), ('at', 'IN'), ('midnight', 'NN'), ('They', 'PRP'), ('are', 'VBP'), ('also',
'RB'), ('happening', 'VBG'), ('in', 'IN'), ('the', 'DT'), ('pop', 'NN'), ('up', 'RP'), ('shop', 'NN'), ('in',
'IN'), ('Dundrum', 'NNP'), ('Town', 'NNP'), ('Centre', 'NNP'), ('too', 'RB')]
```

Moreover, Named Entity Recognizers refers to the classification of unstructured text data in line with pre-defined named entities such as person names, locations, time and quantities. Leon Derczynski et al. state, named entity recognition has achieved 90% accuracy generally on more extended texts, however, it only has 30% - 50% accuracy on microblogs [63]. In other words, it remains challenging to apply named entity recognition to microblogs. In the study, fashion microblogs are analyzed through named entity recognition in NLTK. As seen in the following instance, NER classifies text data of fashion microblogs into general categories, which are insignificant for understanding the content marketing in fashion microblogging. Hence, the study re-trains pre-defined named entities to further identify distinctive entities in the fashion industry, such as brands and products.

NLTK NER:

```
>>> nltk.chunk.ne_chunk(nltk.pos_tag (nltk.word_tokenize (text)))
>>> Tree ('S', [('Yes', 'VB'), ('Our', 'PRP$'), ('winter', 'NN'), ('sales', 'NNS'), ('on', 'IN'),
('pocobypippa.com', 'NN'), ('and', 'CC'), ('pippacollection.com', 'NN'), ('are', 'VBP'), ('ending',
'VBG'), ('tomorrow', 'NN'), ('at', 'IN'), ('midnight', 'NN'), ('They', 'PRP'), ('are', 'VBP'), ('also',
'RB'), ('happening', 'VBG'), ('in', 'IN'), ('the', 'DT'), ('pop', 'NN'), ('up', 'RP'), ('shop', 'NN'), ('in',
'IN'), Tree ('GPE', [('Dundrum', 'NNP')]), ('Town', 'NNP'), ('Centre', 'NNP'), ('too', 'RB')])
```

3.2.2. Semantic Analysis

The second stage is the Semantic Analysis, which consists of Geo-Location Detection, Opinion Mining, Topic Detection, and Automatic Summarization. It concentrates on the discussion of topic

detection and classification from the perspective of text mining analysis. Referring to Ismail Hmeidi et al., text classification, usually referring to text categorization, is defined as a process of “classifying an unstructured text document in its desired category(s) depending on its contents” [64]. Among methods of text classification, automatic keyword extraction is an important research direction in text mining and natural language processing because it enables us to summarize the entire document [65,66]. Therefore, the microblogs are categorized on the basis of keyword classification in the study. For instance, the keywords of Pippa’s microblog mentioned above can be further extracted by NLTK (see Figure 1). The program code “`nltk.FreqDist(nltk.tokenize.word_tokenize(text))`” reveals the word frequency of each word in the microblogs, and shows words from the most to the least frequent. The “`nltk.FreqDist(nltk.tokenize.word_tokenize(text)).freq(' ')`” presents the frequency of a specific word in the microblog. For instance, the frequency of the word “sales” in the microblog is 0.037. Finally, the program code “`nltk.FreqDist(nltk.tokenize.word_tokenize(text)).plot()`” allows researchers to demonstrate the distribution of word frequency in the microblog through line charts. Please see the following details.

```
>>> nltk.FreqDist(nltk.tokenize.word_tokenize(text))
>>> FreqDist({'in': 2, 'are': 2, 'sales': 1, 'tomorrow': 1, 'pop': 1, 'Town': 1, 'Our': 1,
'pocobypippa.com': 1, 'also': 1, 'at': 1, ...})
>>> nltk.FreqDist(nltk.tokenize.word_tokenize(text)).freq('sales')
>>> 0.037037037037037035
>>> nltk.FreqDist(nltk.tokenize.word_tokenize(text)).plot()
>>>
```

Figure 1. Semantic Analysis by NLTK.

3.3. Data Results

As a result, the keywords from microblogs in the study are classified into four groups: brands, products, occasions, and entertainments.

3.3.1. Brands

Since micro-influencers are eager to be the first for spreading the latest news on fashion brands, without doubt, brands are one of the most frequently mentioned words in content marketing. The study finds micro-influencers, Help My Style and Boy Mr K in particular, microblog many brand names in the posts (see Figure 2). The brands are various, ranging from luxury brands (e.g., Gucci, Armani) to affordable brands (e.g., Kenzo, Kiehl's). For Irish microblogging, the study finds that most of the luxury brands are mentioned by market mavens. They attract consumers to notice the styles of luxury brands in content marketing and then recommend affordable products from other brands or online shops. Except for market mavens, other Irish micro-influencers rarely introduce luxury brands. Instead, they market affordable brands directly. By contrast, Chinese micro-influencers hardly ever talk about affordable brands. Luxury brands such as Louis Vuitton are overwhelmed by content marketing. However, both Irish and Chinese micro-influencers prefer to emphasize brands in the capital and bold letters in the microblogging.

3.3.2. Products

The study discovers that micro-influencers (Love Lauren, Just Jordan, Fluff and Frippers, Qiangkouxiaolajiao, Hanhuohuo, and Yang Fan Jame) prefer to review fashion information and give recommendations for products on the basis of their experience to target consumers. Among them, Irish micro-influencers, Just Jordan for example, use words such as look, dress, shoes, and bag frequently in content marketing (see Figure 3a). They tend to introduce a variety of fashion products in the microblogs by selfies, photos, and links. Unlike company marketers' branding, micro-influencers' evidence of using products is more persuasive. Besides, they incline to use positive verbs and adjectives (e.g., best, favorite, love) in the content of marketing. Relatively, Chinese micro-influencers, Qiangkouxiaolajiao for instance, the most frequently used words consist of small, color, dots, wool, knit, silhouette, down, etc. (see Figure 3b). They further reveal the details of fashion products such as fabric (e.g., wool, down), color (grey), and silhouette (e.g., pattern, bottom). The specified information of products guides consumers to understand fashion trends and decide to purchase.

Figure 2. Cont.

Figure 2. The Keywords from (a) Help My Style and (b) Boy Mr K.

3.3.3. Occasions

On the ground of product advice, micro-influencers, especially Irish micro-influencers Anouska, The Style Fairy, Thunder and Threads, Whatshe wears, combine the recommendation with consumers' fashion needs for occasions and enhance consumers' acceptance of marketing messages. Take Whatshe wears for instance. The frequent occasions consist of four seasons, weather, holidays (e.g., Christmas, New Year), and other special occasions (e.g., Irish Payday, Tuesday Shoe day) (see Figure 4a). The micro-influencers give consumers their opinions on what to wear according to different occasions and help to solve consumers' needs. Additionally, another frequent occasion is the location. As seen in Figure 4b, Thunder and Threads, for instance, prefer to microblog fashion according to various places like Dublin, London, and England. In such a case, the micro-influencers connect fashion marketing to tourism and advise proper fashion styles for various places.

3.3.4. Entertainments

Compared with Irish microblogging occasions, Chinese micro-influencers (Mr Kira, Miss Shopping Li, Peter Xu, Chrison, Gogoboi, and Shiliupobaogao) focus on gossiping entertainment news to engage with fashion consumers online. For instance, Figure 5 presents Peter Xu's fashion microblogs contain many names of celebrities like Wu Yifan and Liu Yifei. It indicates that micro-influencers incline to use celebrities' fashion styles to influence consumers in the social network. Considering keyword results from other Chinese fashion micro-influencers, the entertainments in the microblog marketing are summarized as three categories: (1) Celebrities. The micro-influencers introduce celebrities' latest fashion styles and criticize them in the microblogs. (2) Popular movies and TV dramas. The micro-influencers analyze the fashion styles of main characters, find similar products, and persuade consumers to buy them. For example, 1345 pieces of the red lipstick from Yves Saint Laurent in the movie named My Love from the Star were sold in 30 days on account of its outstanding role for leading actor and actress on the film [67]. (3) Hot issues. The micro-influencers try their best to link content marketing with hot issues to arouse consumers' attention and maintain relations with them actively.

(a)

(b)

Figure 3. The Keywords from (a) Just Jordan and (b) Qiangkouxiaolajiao.

(a)

(b)

Figure 4. The Keywords from (a) Whatsheweares and (b) Thunder and Threads.

Figure 5. The Keywords from Peter Xu.

4. The Development of Digital Artifact

For the sustainable development of research results, the study presents the findings in the form of a digital artifact. On the whole, there are three steps for developing this digital artifact: (1) Database layer, (2) Business layer, (3) Presentation Layer. The detailed design is outlined in Figure 6.

Figure 6. The Design of Digital Artifact.

The first step is to build the Database Layer. This step is further divided into two stages: (1) Train classifiers based on testing text data, (2) Automatically classify unknown text data according to pre-defined classifiers. On the ground of data results, the study trains four classifiers—brands, products, occasions, and entertainments—for the digital artifact. Each classifier consists of a wordlist of keywords extracted from the microblogs in the study. For example, keywords such as Louis Vuitton, fishnet, this week and Gigi Hadid can be re-trained as below. Firstly, the NLTK library is imported to programming. Then, we define Louis Vuitton as one tokenization named brand through POS. Likewise, we apply the same programming to define fishnet, this week and Gigi Hadid for re-training POS taggers. After the re-training, the words are tagged as brand, product, occasion, and entertainment automatically when these words occur in the texts. It can extract not only fashion-related words like fishnet and Topshop but also phrases like Louis Vuitton. The collocation of fashion words and phrases help understand what kind of microblogging content affects fashion consumers significantly.

Re-trained POS Taggers:

```

>>> import nltk
>>> pos={'Louis Vuitton': 'brand'}
>>> pos['Louis Vuitton']
>>> 'brand'
>>> pos={'fishnet': 'product'}
>>> pos['fishnet']
>>> 'product'
>>> pos={'this week': 'occasion'}
>>> pos['this week']
>>> 'occasion'
>>> pos={'Gigi Hadid': 'entertainment'}
>>> pos['Gigi Hadid']
>>> 'entertainment'

```

Subsequently, it comes up with the second step—the Business Layer, which involves two steps (see Figure 7). Firstly, it extracts keywords from random fashion microblogs provided by the users of digital artifact automatically by means of text mining. Then, the extracted keywords are compared with pre-defined four classifiers in the database layer. Finally, the successful pairing of keywords leads to the third step—the Presentation Layer, which refers to the final digital artifact.

Figure 7. The Business Layer.

The digital artifact is named as Fashion Microblogging Marker, and the web hosting server is Reclaim Hosting. At present, there are two popular web hosting server WordPress and Reclaim Hosting. WordPress (WordPress.com), developed by Matt Mullenweg and Mike Little in 2003, is one of the most natural methods for the creation of free websites and blogs. Till April 2018, it has been utilized by 32.1% of all the websites, which occupies 59.5% of the market share of the content management system [68]. The high adoption of WordPress is caused by its interactive, sharing and social features. In terms of Reclaim Hosting (reclaimhosting.com), it was created by Jim Groom and Tim Owens in 2013. Similar to WordPress, it consists of several free plugins such as Scalar, Omeka, and Drupal, which allow users to add, edit and develop their websites. One of the outstanding characteristics for Reclaim Hosting is to enable users to have their domain names without undesired domain suffixes.

By comparison, this study considers that Reclaim Hosting is a more suitable web hosting server for this research. For one thing, it is attractive to have individual domain names. WordPress allows researchers to have their domain names conditionally. The names should end with suffixes such as WordPress. In order to remove these suffixes, researchers have to pay for their domain names like

Reclaim Hosting. For another, WordPress is included in the Reclaim Hosting as one of the free plugins. Although this WordPress does not have all functions of original WordPress, it has enough functions for researchers to build websites. At the same time, Reclaim Hosting has many other free plugins, which enable researchers to apply them. Moreover, Reclaim Hosting turns out to be more comfortable for users to code the websites as their wishes. As for WordPress, it contains more given templates and themes of websites, which seem to be simple for users to construct the websites. However, it limits the room for personalized websites. On the contrary, Reclaim Hosting gives web developers more freedom to design their websites. As a consequence, this study chooses Reclaim Hosting as the web hosting server for this digital artifact eventually.

The framework of the presentation layer starts with users' typing, copying and pasting the content of microblogs into the artifact. The content is given the value *m*. Then, the artifact extracts keywords from *m* automatically, and the extracted keywords get a new value *k*. We set the value *t* in the artifact previously. Value *t* contains four values—*b*, *p*, *o*, and *e*, which represent the brand, product, occasion and entertainment keywords included in the database layer. Afterward, *k* compares with *t*. If *k* = *b*, *k* is highlighted in green, and the program ends. If not, *k* continues to compare with other values. If *k* = *p*, *k* is highlighted in blue, and the program ends. If *k* = *o*, *k* is highlighted in yellow, and the program ends. If *k* = *e*, *k* is highlighted in pink, and the program ends. If *k* ≠ *e*, then the artifact outputs *m* simply and goes back to the input *m*. It indicates that bloggers need to revise the content in the posts. The detailed framework of the presentation layer is demonstrated in Figure 8.

Consequently, Figure 9 presents the final digital artifact. It is hosted by Reclaim Hosting. The interface consists of three main parts. At the top, a random example of micro-influencers' content is provided for users to test. Next, it provides warm tips, which list some common keywords in the study for reference. Subsequently, it comes up with notes to explain the meaning of results. Eventually, it is a textbox for users to input the marketing content, which follows the framework in Figure 8. The final result is shown in the same textbox spontaneously. When all or most of the texts are colored in the textbox, it means that this microblog contains many keywords of fashion microblog marketing. Thus, this microblog can increase consumers' traffic and influence online based on keywords in content marketing and SEO. The marketers who think about employing blogs to target consumers can take advantage of this artifact to draft and revise the content in the blogs, and become influencers among members of the social network through effective content marketing.

According to Thomas E. Epalle, there are five challenges for computer-assisted mining microblog marketing—data extraction and pattern evaluation, social media analysis, natural language processing, opinion mining, and sentiment analysis, and China Internet censorship, and proposes to develop new application-programming interfaces (APIs) due to the lack of open and free applications to mining microblog marketing [69]. Therefore, this study manages to develop an open and free digital artifact to solve the problem. The artifact is designed to help marketers and bloggers model their content for achieving marketing purposes. It has several distinctive characteristics:

(1) Fast processing without a word limit. There is no maximum word for the analysis at one time. Nowadays, the problems of many applications on text mining are characterized by word limitations and difficulties in analyzing big data. For example, ICTCLAS can only analyze a maximum of 3000 words at one time, and Voyant Tools proceeds text analysis slowly when data come from more than ten websites. Hence, this artifact is programmed to process big data efficiently and effectively corresponding to the popularity of social media and the era of big data.

(2) Bilingual. At present, applications do not support text analysis in multi-languages at the same time, such as NLTK and Stanford NLP toolkit. As far as I know, this artifact is the first application to support English and Chinese text analysis at the same time. It can process not only English text data but also Chinese text data in content marketing. More importantly, it can analyze English and Chinese text data at the same time. Chinese micro-influencers like Boy Mr K and Han Huohuo mix English words with Chinese texts in the microblogs. Thus, this artifact improves the analysis accuracy of multilingual content in microblog marketing. Besides, the interface is written in English and Chinese

as well, which allows more scholars to use this artifact, conduct relevant research and increase the influence of this study internationally.

(3) Enhancing the accuracy of analyzing content in marketing. It improves text analysis of marketing on social media, in particular, fashion marketing. For instance, it is confusing when the word frequencies of brown and Thomas are high in the analysis. It is uncertain about summarizing whether brown is a new trend of color in fashion, Thomas is a famous fashion figure at present, or Brown Thomas is a popular shopping center in Ireland. As current applications fail to recognize fashion-related words in the text mining analysis, scholars tend to solve the problem manually. Generally, it takes a long time, especially big data. Therefore, this artifact reduces researchers' time and effort. On the ground of the keyword database in the study, it recognizes more fashion-related texts and enhances the accuracy of POS and NER in fashion marketing.

Figure 8. The Framework of the Presentation Layer.

Fashion Microblogging Marker

Example 1 例1: The first of my big announcements this week I'm thrilled to announce that I'm collaborating with everyone's favourite retail giant Penneys on an EXCLUSIVE lash collection. As you saw from my teasers over on SOSU by Suzanne Jackson, I've designed 3 deluxe lashes that can only be found in Penneys/Primark. Price: €6 Kylie Belle Holly I'll be doing an in store launch later this week in Swords, more details to follow! Who's excited?

Warm Tips Missguided Jo Malone this week white shirts little black dress favourite winter Cork Black
Friday Penneys Gigi Hadid Olivia Palermo

Notes 注释: Texts in the fashion microblogs are marked as **BRAND** **PRODUCT** **OCCASION** **ENTERTAINMENT**

The first of my big announcements **this week** I'm thrilled to announce that I'm collaborating with everyone's **favourite** retail giant **Penneys** on an EXCLUSIVE **lash collection**. As you saw from my teasers over on **SOSU** by **Suzanne Jackson**, I've designed 3 deluxe **lashes** that can only be found in **Penneys/Primark**. Price: €6 Kylie Belle Holly I'll be doing an **in store** launch later **this week** in **Swords**, more **details** to follow! Who's excited?

1

Figure 9. The Presentation of Digital Artifact.

5. Conclusions

In summary, this study proposes to design a digital artifact for demonstrating findings of marketing research through three stages—database layer, business layer, and presentation layer. The study investigates fashion microblogs from top Irish and Chinese micro-influencers by means of text mining analysis. As a result, it figures out four categories of keywords in micro-influencers' content marketing on social media—Brands, Products, Occasions, and Entertainment. The keywords of the four categories are used to re-train classifiers in the database layer. Afterward, the business layer is programmed to extract keywords from fashion microblogs, and analyze them employing NLTK. Eventually, the artifact is shown in the presentation layer, which consists of a scientific framework and visualization. The proposed digital artifact not only contributes to model bloggers' marketing content in fashion microblogging but also represents the first case for disseminating results of studies on content marketing in fashion on the ground of creating digital tools. Consequently, this study enlightens more scholars to apply the development of digital artifacts as an innovative way to present the results of marketing analysis.

5.1. Contributions to Existing Knowledge

Firstly, this paper extends previous studies on analyzing online marketing content employing text mining techniques and programming codes. Humphreys and Wang state that automated text analysis is a useful tool for discovering patterns in text that marketing scholars are unable to detect [70]. Matthew K. Gold further proposes that researchers need to explore how the creation and deployment of such a tool perform distinct functions to achieve the maximal effect [71]. Hence, it contributes to the literature on how to develop the tool of text analysis in content marketing.

Furthermore, the key contribution is the demonstration of findings in the form of a digital artifact. With the advance of technologies, scholars argue the challenges that technologies bring to marketing research. The challenges include consumer privacy, societal equality, and individual wellbeing. For example, Lancelot Miltgen et al. discuss how privacy concerns can affect consumers' intention to accept IT innovation, and Krafft et al. investigate permission marketing and consumers' related privacy concerns [72,73]. In line with Christian Rauch, "While surely, technology is also among the culprits of many of our problems, it is clear that technological development and innovation will be at the heart of moving towards a sustainable future" [74]. For instance, artificial intelligence benefits to understand

markets more clearly along with big data. Michael Chui et al. contend that using artificial intelligence help achieve sustainable development [75]. Supported by Christian Rauch, the application of different machine learning techniques like training data in artificial intelligence provides solutions to solve sustainable problems.

Therefore, the paper gives more insights for marketing researchers to achieve content marketing in the fashion microblogging. It indicates that they could develop digital artifacts on the ground of natural language processing for the sustainable development of marketing research. The digital artifact is an eco-friendly and sustainable form of presenting research results. The pulp and paper industry is regarded as one of the most polluting industries in the world as a result of the printing press, paper mills, paper consumption and waste [76]. The paper pollution further leads to air pollution, water pollution, climate change, wood waste and sludge [77]. For this reason, digital artifact lies in the World Wide Web, which reduces paper pollution and waste caused by traditional presentations of research results such as books and journals. Moreover, digital artifact allows users to develop marketing findings sustainably. The research data can be used, shared, and reused by other researchers through digital artifacts. The researchers are able to edit and revise findings in the digital artifacts constantly, which contributes to the sustainable influence of research results in the academy. Thus, scholars could consider the creation of digital artifacts as a presentational form of research results.

5.2. Implications for Practice

The proposed digital artifact turns theoretical findings into the practical application, which provides more significant implications of research results. Referring to Arokia R. Terrance et al., the artifact can help enhance the rank of microblogs in the search engines as a result of keyword analysis [18]. It contributes to evaluate and predict the marketing effect of fashion microblogs in the social network. Bloggers and marketers can be notified of the influence of content marketing by the artifact in advance, and then revise fashion microblogs according to analysis results. The artifact assists bloggers and marketers not only to solve the problem of oversaturation in blogging nowadays but also to convince consumers by means of smart content marketing and personalized microblogging. Also, marketing researchers and text miners can apply the artifact to understand and analyze fashion microblogs in the relevant research projects from the perspective of text mining. The artifact improves English and Chinese tokenization of fashion-related words and phrases in the text mining analysis. Last but not least, the bilingual artifact enables international researchers to practice handily.

5.3. Limitations and Future Research Directions

This paper is subject to some limitations, which enlightens new directions for further studies. The database layer of the digital artifact is built on the ground of research results from specific contexts, namely, microblogs of Irish and Chinese micro-influencers. Arlesa J. Shephard et al. compare Hispanic and Caucasian fashion leaderships, and prove that “differences exist between ethnic groups, between genders, and between genders within the Hispanic group and among Hispanic fashion leaders” [56]. The database layer may differ if research data are from fashion micro-influencers of other countries. Future research could consider different contexts of micro-influencers such as Hispanic and Caucasian fashion leaders. In addition, digital artifact employs the keyword analysis to develop the business layer because it is certified to increase the influence of content marketing [14–17]. However, other algorithms like Latent Dirichlet Allocation, Probabilistic Latent Semantic Analysis, Biterm Topic Model and K-means clustering algorithm can also analyze the content of microblogging [78]. Future studies can develop the business layer of the artifact based on various methods.

Author Contributions: Writing—original draft preparation, Z.S.; writing—review and editing, supervision, A.d.l.G.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Belk, W.R.; Llamas, R. Living in a digital world. In *The Routledge Companion to Digital Consumption*; Belk, W.R., Llamas, R., Eds.; Routledge, Taylor and Francis Group: London, UK, 2013; pp. 3–12.
2. Hanser, A.; Li, J.L.C. Opting out? gated consumption, infant formula and China's affluent urban consumers. *China J.* **2015**, *74*, 110–128. [CrossRef]
3. Statista. Fashion Ireland. Available online: <https://www.statista.com/outlook/244/117/fashion/china?currency=eur> (accessed on 28 April 2019).
4. Kim, J.A.; Ko, E. Do social media marketing activities enhance customer equity? an empirical study of luxury fashion brand. *J. Bus.* **2012**, *65*, 1480–1486. [CrossRef]
5. Lenhart, A.; Purcell, K.; Smith, A.; Zickuhr, K. Social Media & Mobile Internet Use among Teens and Young Adults. *Pew Research Center*. Available online: <https://www.pewresearch.org/internet/2010/02/03/social-media-and-young-adults/> (accessed on 19 November 2019).
6. Arora, A.; Bansal, S.; Kandpal, C.; Aswani, R.; Dwivedi, Y. Measuring social media influencer index-insights from facebook, twitter and instagram. *J. Retail. Consum. Serv.* **2019**, *49*, 86–101. [CrossRef]
7. Jefferson, S.; Tanton, S. *Valuable Content Marketing: How to Make Quality Content Your Key to Success*; Kogan Page: London, UK, 2015.
8. Burley, D.; Silver, N. The impact of 'smart content' and metadata from creation to distribution. *J. Digit. Media Manag.* **2017**, *5*, 326–335.
9. Kozinets, R.V. *Netnography: The Marketer's Secret Weapon*; NetBase: Santa Clara, CA, USA, 2010.
10. Araujo, T.; Neijens, P.; Vliegenthart, R. Getting the word out on Twitter: The role of influentials, information brokers and strong ties in building word-of-mouth for brands. *Int. J. Advert.* **2016**, *36*, 496–513. [CrossRef]
11. Barker, M.; Barker, D.; Bormann, N.; Neher, K. *Social Media Marketing: A Strategic Approach*; South-Western, Cengage Learning: Boston, MA, USA, 2013.
12. Hughes, C.; Swaminathan, V.; Brooks, G. Driving brand engagement through online social influencers: An empirical investigation of sponsored blogging campaigns. *J. Mark.* **2019**, *83*, 78–96. [CrossRef]
13. Ramirez-de-la-Rosa, G.; Villatoro-Tello, E.; Jimenez-Salazar, H.; Sanchez-Sanchez, C. Towards automatic detection of user influence in Twitter by means of stylistic and behavioural features. In *Human-Inspired Computing and Its Applications*; Gelbukh, A., Espinoza, G.C., Galicia-Haro, S.N., Eds.; MICAI: Tuxtla Gutierrez, Mexico, 2014; Volume 8856, pp. 245–256.
14. Moldovan, S.; Muller, E.; Richter, Y.; Tov, E.Y. Opinion leadership in small groups. *Int. J. Res. Mark.* **2017**, *34*, 536–552. [CrossRef]
15. Dubois, E.; Gaffney, D. The multiple facets of influence: Identifying political influentials and opinion leaders on Twitter. *Am. Behav. Sci.* **2014**, *58*, 1260–1277. [CrossRef]
16. Tuarob, S.; Tucker, C.S. Automated discovery of lead users and latent product features by mining large scale social media networks. *J. Mech. Des.* **2015**, *137*, 1–11. [CrossRef]
17. Macskassy, S.A. Contextual linking behavior of bloggers: Leveraging text mining to enable topic-based analysis. *Soc. Netw. Anal. Min.* **2011**, *1*, 355–375. [CrossRef]
18. Basili, C.; Lanzillo, L. Research quality criteria in the evaluation of books. In *The Evaluation of Research in Social Sciences and Humanities. Lessons from the Italian Experience*; Bonaccorsi, A., Ed.; Springer: Berlin, Germany, 2018; pp. 159–184.
19. Engels, T.; Starcic, A.I.; Kulczycki, E.; Pölönen, J. Are book publications disappearing from scholarly communication in the social sciences and humanities? *Cent. Sci. Technol.* **2018**, *70*, 774–780. [CrossRef]
20. Greenwood, S.; Perrin, A.; Duggan, M. Social Media Update 2016. *Pew Research Center: Internet & Technology*. Available online: <http://www.pewinternet.org/2016/11/11/social-media-update-2016/> (accessed on 28 November 2017).
21. Alperin, J.P.; Gomez, C.J.; Haustein, S. Identifying diffusion patterns of research articles on Twitter: A case study of online engagement with open access articles. *Public Underst. Sci.* **2019**, *28*, 2–18. [CrossRef] [PubMed]
22. Sugimoto, C.R.; Work, S.; Larivière, V.; Haustein, S. Scholarly use of social media and altimetric: A review of the literature. *J. Assoc. Inf. Sci. Technol.* **2017**, *68*, 2037–2062. [CrossRef]
23. Sterne, J. *Artificial Intelligence for Marketing: Practical Applications*; John Wiley and Sons: Hoboken, NJ, USA, 2017.

24. Davenport, T.; Guha, A.; Grewal, D.; Bressgott, T. How artificial intelligence will change the future of marketing. *J. Acad. Mark. Sci.* **2019**, 1–19. [CrossRef]
25. Hanlon, A. *Digital Marketing: Strategic Planning & Integration*; Sage Publications Ltd.: London, UK, 2019.
26. Hall, J. 7 Marketing Trends to Budget for in 2019. Available online: <https://www.forbes.com/sites/johnhall/2018/11/11/7-marketing-trends-to-budget-for-in-2019/#d43ee82511c8> (accessed on 11 November 2018).
27. DeMers, J. 5 Ways A Content Marketing Campaign Can Backfire. Available online: <https://www.forbes.com/sites/jaysondemers/2015/10/09/5-ways-a-content-marketing-campaign-can-backfire/#aab75065e0b7> (accessed on 9 October 2015).
28. Forbes Agency Council. 10 Marketing Tools and Tactics That Will Shake Up the Industry in 2019. Available online: <https://www.forbes.com/sites/forbesagencycouncil/2019/01/25/10-marketing-tools-and-tactics-that-will-shake-up-the-industry-in-2019/#2fe47f803412> (accessed on 25 January 2019).
29. Lieb, R. *Content Marketing*; Que Publishing: Indianapolis, Indiana, 2012.
30. Jerath, K.; Liya, M.; Park, Y.H. Consumer click behaviour at a search engine: The role of keyword popularity. *Keller Cent. Res. Rep.* **2015**, 8, 21–24.
31. Simonov, A.; Nosko, C. Competition and crowd-out for brand keywords in sponsored search. *Mark. Sci.* **2018**, 37, 177–331. [CrossRef]
32. Lu, S.J.; Yang, S. Investigating the spillover effect of keyword market entry in sponsored search advertising. *Mark. Sci.* **2017**, 36, 813–1017. [CrossRef]
33. Terrance, A.R.; Shrivastava, S.; Kumari, A. Importance of search engine marketing in the digital world. *Technol. Knowl. Manag.* **2018**, 14, 155–158.
34. Harris, J. Road Map to Success: Content Marketing Strategy Essentials. Available online: <https://contentmarketinginstitute.com/2018/02/content-strategy-essentials/> (accessed on 20 February 2018).
35. Tekin, M.; Etlioglu, M.; Koyuncuoglu, O. Data mining in digital marketing. In Proceedings of the International Symposium for Production Research 2018; Springer: Berlin, Germany, 2018; pp. 44–61.
36. Bello, O.G.; Jung, J.J.; Camacho, D. Social big data: Recent achievements and new challenges. *Inf. Fusion* **2016**, 28, 45–59. [CrossRef]
37. Zhang, K.; Stoffel, A.; Behrisch, M.; Mittelstadt, S.; Schreck, T.; Pompl, R.; Weber, S.; Last, H.; Keim, D. Visual analytics for the big data era—A comparative review of state-of-the-art commercial systems. In Proceedings of the 2012 IEEE Conference on Visual Analytics Science and Technology (VAST), Seattle, WA, USA, 14–19 October 2012; pp. 173–182.
38. Wang, W.; Chen, L.; Thirunarayan, K.P.; Sheth, A. Harnessing twitter ‘big data’ for automatic emotion identification. In Proceedings of the 2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing, Amsterdam, The Netherlands, 3–5 September 2012; pp. 1–6.
39. Talib, R.; Hanif, M.K.; Ayesha, S.; Fatima, F. Text mining: Techniques, applications and issues. *Int. J. Adv. Comput. Sci. Appl.* **2016**, 7, 414–418. [CrossRef]
40. Weiss, S.M.; Indurkha, N.; Zhang, T.; Damerau, F.J. *Text Mining: Predictive Methods for Analyzing Unstructured Information*; Springer Science & Business Media: New York, NY, USA, 2010.
41. Muninger, M.; Hammedi, W.; Mahr, D. The value of social media for innovation: A capability perspective. *J. Bus. Res.* **2019**, 95, 116–127. [CrossRef]
42. Holzinger, A.; Schantl, J.; Schroettner, M. Biomedical text mining: State-of-the-art, open problems and future challenges. In *Knowledge Discovery and Data Mining*; Holzinger, A., Jurisica, I., Eds.; Springer-Verlag Berlin Heidelberg: Berlin, Germany, 2014; pp. 271–300.
43. Glowacki, E.M.; Glowacki, J.B.; Wilcox, G.B. A text-mining analysis of the public’s reactions to the opioid crisis. *J. Subst. Abus.* **2017**, 7, 1–5. [CrossRef] [PubMed]
44. Bae, J.H.; Song, J.E.; Song, M. Analysis of Twitter for 2012 South Korea presidential election by text mining techniques. *J. Intell. Inf. Syst.* **2013**, 19, 141–156.
45. Bowers, A.J.; Chen, J.J. Ask and ye shall receive? Automated text mining of michigan capital facility finance bond election proposals to identify which topics are associated with bond passage and voter turnout. *J. Educ. Financ.* **2015**, 41, 164–196.
46. Hosoi, H.; Yamagata, T.; Ikarashi, Y. Visualization of special features in “The Tale of Genji” by text mining and correspondence analysis with clustering. *J. Flow Control Meas. Vis.* **2014**, 2, 1–6. [CrossRef]

47. Westergaard, D.; Staerfeldt, H.H.; Tonsberg, C.; Jensen, L.J.; Brunak, S. Text mining of 15 million full-text scientific articles. *Biorxiv* **2017**, *11*, 1–28.
48. Hung, J. Trends of e-learning research from 2000 to 2008: Use of text mining and bibliometrics. *Br. J. Educ. Technol.* **2012**, *43*, 5–16. [\[CrossRef\]](#)
49. Yasuhara, T.; Sone, T.; Konishi, M.; Kushihata, M.; Nishikawa, T.; Yamamoto, Y.; Kurio, W.; Kohno, T. Studies using text mining on the differences in learning effects between the KJ and world café method as learning strategies. *J. Pharm. Soc. Jpn.* **2015**, *135*, 753–759. [\[CrossRef\]](#)
50. Mostafa, M.M. More than words: Social networks' text mining for consumer brand sentiments. *Expert Syst. Appl.* **2013**, *40*, 4241–4251. [\[CrossRef\]](#)
51. Culotta, A.; Cutler, J. Mining brand perceptions from twitter social networks. *Mark. Sci.* **2016**, *35*, 1–39. [\[CrossRef\]](#)
52. Owyang, J.; Lovett, J. Social Marketing Analytics A New Framework for Measuring Results in Social Media. Web Analytics Demystified and Altimeter Group. Available online: <https://www.slideshare.net/jlovett/social-marketing-analytics> (accessed on 3 October 2019).
53. Brown, D.; Fiorella, S. *Influence Marketing: How to Create, Manage, and Measure Brand Influencers in Social Media Marketing*; Que Publishing: Upper Saddle River, NJ, USA, 2013.
54. Boerman, S.C. The effects of the standardized instagram disclosure for micro- and meso-influencers. *Comput. Hum. Behav.* **2020**, *103*, 199–207. [\[CrossRef\]](#)
55. Ersun, A.N.; Yildirim, F. Consumer involvement and brand sensitivity of university students in their choice of fashion. *Marmara Univ. J. Fac. Econ. Administ* **2010**, *28*, 313–333.
56. Shephard, A.; Kinley, T.R.; Josiam, B.M. Fashion leadership, shopping enjoyment, and gender: Hispanic versus, Caucasian consumers' shopping preferences. *J. Retail. Consum. Serv.* **2014**, *21*, 277–283. [\[CrossRef\]](#)
57. Sarathy, P.; Patro, S. The role of opinion leaders in high-involvement purchases: An empirical investigation. *S. Asian J. Manag.* **2013**, *20*, 11–53.
58. Yang, H. Market mavens in social media: Examining young Chinese consumers' viral marketing attitude, eWOM motive, and behaviour. *J. Asia-Pac. Bus.* **2013**, *14*, 154–178. [\[CrossRef\]](#)
59. Allahyari, M.; Pouriyeh, S.; Assefi, M.; Safaei, S.; Trippe, E.D.; Gutierrez, J.B.; Kochut, K. A brief survey of text mining: Classification, clustering and extraction techniques. *arXiv* **2017**, arXiv:1707.02919 2017, 1–13.
60. Farzindar, A.; Inkpen, D. Natural language processing for social media. *Synth. Lect. Hum. Lang. Technol.* **2017**, *10*, 1–195. [\[CrossRef\]](#)
61. Bird, S. *Natural Language Processing with Python*; O'Reilly Media. H: Sebastopol, CA, USA, 2010.
62. Gaikwad, V.S.; Chaugule, A.; Patil, P. Text mining methods and techniques. *Int. J. Comput. Appl.* **2014**, *85*, 42–45.
63. Dercynski, L.; Maynard, D.; Rizzo, G.; Erp, M.; Gorrell, G.; Troncy, R.; Petrak, J.; Bontcheva, K. Analysis of named entity recognition and linking for tweets. *Inf. Process. Manag.* **2015**, *51*, 32–49. [\[CrossRef\]](#)
64. Hmeidi, I.; Al-Ayyoub, M.; Abdulla, N.; Almodawar, A.A. Automatic Arabic text categorization: A comprehensive comparative study. *J. Inf. Sci.* **2014**, *41*, 1–11.
65. Wu, X.; Du, Z.K.; Guo, Y.K. A visual attention-based keyword extraction for document classification. *Multimed. Tools Appl.* **2018**, *77*, 25355–25367. [\[CrossRef\]](#)
66. Onan, A.; Korukoglu, S.; Bulut, H. Ensemble of keyword extraction methods and classifiers in text classification. *Expert Syst. Appl.* **2016**, *57*, 232–247. [\[CrossRef\]](#)
67. Chen, Y.; Liu, X.Y. Recognizing opinion leaders based on social network analysis. *J. Inf. Sci.* **2015**, *33*, 13–20.
68. W3Techs. Usage of Content Management Systems for Websites. Available online: https://w3techs.com/technologies/overview/content_management/all/ (accessed on 23 April 2018).
69. Epalle, T.E. Mobile social media mining challenges overview: A case study of wechat. *Int. J. Comput. Technol. Appl.* **2015**, *6*, 347–351.
70. Humphreys, A.; Jen-Hui Wang, R. Automated text analysis for consumer research. *J. Consum. Res.* **2017**, *44*, 1274–1306. [\[CrossRef\]](#)
71. Klein, L.F.; Gold, M.K. *Debates in the Digital Humanities 2016*; University of Minnesota Press: Minneapolis, MN, USA, 2016.
72. Miltgen, C.L.; Henseler, J.; Gelhard, C.; Popovič, A. Introducing new products that affect consumer privacy: A mediation model. *J. Bus. Res.* **2016**, *69*, 4659–4666. [\[CrossRef\]](#)

73. Krafft, M.; Arden, C.M.; Verhoef, P.C. Permission marketing and privacy concerns—why do customers (not) grant permissions? *J. Interact. Mark.* **2017**, *39*, 39–54. [CrossRef]
74. Rauch, C. AI for Good—How Artificial Intelligence Can Help Sustainable Development. Available online: <https://medium.com/@C8215/ai-for-good-how-artificial-intelligence-can-help-sustainable-development-58b47d1c289a> (accessed on 4 November 2019).
75. Chui, M.; Chung, R.; Heteren, A. Using AI to Help Achieve Sustainable Development Goals. Available online: https://www.undp.org/content/undp/en/home/blog/2019/Using_AI_to_help_achieve_Sustainable_Development_Goals.html (accessed on 4 October 2019).
76. Ince, B.K.; Cetecioglu, Z.; Ince, O. Pollution Prevention in the Pulp and Paper Industries. *Environmental Management in Practice*, 2011. Available online: <https://www.intechopen.com/books/environmental-management-in-practice/pollution-prevention-in-the-pulp-and-paper-industries> (accessed on 4 October 2019).
77. Gavrilescu, D.; Puitel, A.C.; Dutuc, G.; Craciun, G. Environmental impact of pulp and paper mills. *Environ. Eng. Manag. J.* **2012**, *11*, 81–85. [CrossRef]
78. Li, W.J.; Feng, Y.M.; Li, D.J.; Yu, Z.T. Micro-blog topic detection method based on BTM topic model and K-means clustering algorithm. *Autom. Control Comput. Sci.* **2016**, *50*, 271–277. [CrossRef]

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).