1-4-2021

The state of #digitalentrepreneurship: A big data Leximancer analysis of social media activity

Violetta Wilk
Edith Cowan University

Helen Cripps
Edith Cowan University

Alexandru Capatina

Adrian Micu

Angela-Eliza Micu

The state of #digitalentrepreneurship: A big data Leximancer analysis of social media activity

Violetta WILK*, Edith Cowan University, Australia, v.wilk@ecu.edu.au, ORCID: 0000-0001-7990-769X

Helen CRIPPS, Edith Cowan University, Australia, h.cripps@ecu.edu.au

Alexandru CAPATINA*, Dunarea de Jos University of Galati, Romania, acapatana@ugal.ro

Adrian MICU, Dunarea de Jos University of Galati, Romania, adrian.micu@ugal.ro

Angela-Eliza MICU, Ovidius University of Constanta, Romania, angelaelizamicu@yahoo.com

*Corresponding authors

Abstract

Purpose: The purpose of this paper is to better understand the state of digital entrepreneurship evident in social media narrative. This paper examined online sentiment, key themes and patterns evident in social media activity about digital entrepreneurship. It provides a snapshot-in-time, visual-first perspective on social media user-generated-content (UGC) to better understand the topic of digital entrepreneurship.

Design/methodology/approach: Global data consisting of 31,017 publicly available UGC which used the #digitalentrepreneurship (hashtag) and the keywords 'digital entrepreneurship' were collected. A computer assisted qualitative data analysis software (CAQDAS), Leximancer, was used for an automated text-mining analysis.

Findings: There is positive online sentiment surrounding digital entrepreneurship technology, ecosystem and industry, and one which promotes women transformation of digital entrepreneurship globally. Negative sentiment pointed out that future development and support of youth in digital entrepreneurship is needed. Digital entrepreneurs were identified as needing to focus on strategy, leadership, management, and social media platforms.

Research implications/limitations: A comprehensive perspective on the state of digital entrepreneurship in online UGC is provided. Insights into the challenges, issues, changes, success stories and key topics in digital entrepreneurship are highlighted. Future research is encouraged to adopt longitudinal and quantitative approaches, to provide further insights into the evolution of digital entrepreneurship.

Originality/value: The paper contributes to the entrepreneurship literature by applying the Social Exchange Theory and the Social Media User Engagement Framework to better understand social media activity around digital entrepreneurship. The findings show that there

are real challenges and issues to overcome but there are also changes occurring in digital entrepreneurship and social media users are keen to share and learn from digital entrepreneurship success stories.

Keywords: digital entrepreneurship, big data, user-generated content, sentiment analysis, Leximancer, Social Exchange Theory.

Declarations:

Introduction

Digital entrepreneurship facilitates the processes of strategic knowledge acquisition and transfer through the use of web-based technologies which are capable of initiating effective ways of doing business (Rogers 2016). Digital entrepreneurship creates new ventures and transforms existing businesses by developing novel digital technologies and/or the novel usage of digital technology (European Commission 2015). People interested in developing entrepreneurial initiatives based on innovative business models have access to a wide range of information shared through user-generated content (UGC) on social media networks. These commonly fall under the broad umbrella of digital entrepreneurship (Richter, Kraus, Brem, Durst and Giselbrecht et al. 2017). Despite UGC about digital entrepreneurship being in abundance on social media networks, it is surprisingly an under-researched area.

In light of the paucity of research on entrepreneurship in the digital age, definitions fall on a continuum between the extremes of pure digital entrepreneurship and pure traditional entrepreneurship (Nzembayie, Buckley and Cooney 2019). This presents a challenge for a qualitative approach using social media analytics and researchers, such as Nambisan (2017), have highlighted the need for methodologies that reflect the contextual analysis of social media content that support digital entrepreneurial initiatives. Despite the important role which social media plays in driving the proliferation of digital entrepreneurship and in shaping entrepreneurial ideas, social media content analysis that uncovers themes based on the frequencies and the strength of co-occurrences of keywords within a large database of social media posts, remains an under-researched area in digital entrepreneurship (Lai and To 2015). As digital entrepreneurship has become a topic of interest for many people across the world, it is important to understand how social media data, such as UGC, can be harvested and analysed.

Such an approach can provide valuable ideas which could inform effective strategies for approaching online market niches and for creating start-ups.

In an attempt to close this research gap, this exploratory research applied text-mining, thematic analysis of big data from social media UGC about digital entrepreneurship, to explore how social media users discuss #digitalentrepreneurship in their UGC. The study contributes towards current understanding of the challenges, issues and 'hot' topics in digital entrepreneurship and seeks to answer the following key questions:

- Which key topics (themes) are evident in UGC about digital entrepreneurship?
- What patterns reflecting sentiment valence are found within UGC about digital entrepreneurship?
- Do the key topics (themes) differ by sentiment valence (positive, negative, neutral)?
- What thematic patterns, specifically aligned with comment, views, following and followers counts, are found in UGC about digital entrepreneurship?

Social Exchange Theory (Blau 1968; Emerson 1976) and Social Media User Engagement Framework (Barger, Peltier and Schultz 2016) were applied in this research. The social media activity related to digital entrepreneurship, specifically the exchange which occurs through comments, views, following and followers, was analysed. We start from the premise that the number of followers an author of social media post has (follower count) may be associated with the extent of social media activity (comments and views count) about digital entrepreneurship.

Social Exchange Theory posits social behaviour as an exchange which may result in tangible and intangible outcomes, such that one's actions are contingent on rewarding reactions from others (Emerson 1976). In the case of social media UGC, social media users might expect the intangible benefits associated with UGC about digital entrepreneurship to play a significant

role in affecting individuals' behaviour regarding entrepreneurial initiatives, as the exchange of tangible benefits in social interactions is atypical (Liu, Min, Zhai and Smyth 2016)..

Further, the Social Media User Engagement Framework might provide a supplementary lens for this study. This framework consists of a set of measurable actions that people take on social media in response to a specific topic, in this case #digitalentrepreneurship, such as: reacting to content (e.g., Likes, hearts, 1 to 5 star ratings), commenting on content (e.g., Facebook comments, Twitter replies), sharing content with others (e.g. Twitter retweets), and posting UGC (e.g. Twitter posts) (Barger, Peltier and Schultz 2016). Our research considers all UGC that form the online narrative about digital entrepreneurship over a specific period of time; hence, our research considers the resulting exchange and the engagement in online discussion about digital entrepreneurship.

The next section of our paper offers an outline of theoretical background upon which our research is based. The subsequent section outlines the research method, followed by the results. The paper concludes by highlighting discussion of findings and by presenting implications and limitations of the study, along with avenues for future research.

Theoretical Background

In the digital entrepreneurship literature, a recent study revealed that a major focus has been placed on understanding key success factors that drive social media usage amongst entrepreneurs, and investigated the impact of social media posts on the creation and exchange of user-generated content (UGC) (Olanrewaju, Hossain, Whiteside and Mercieca 2020). Entrepreneurs, organisations and communities interested in digital business can use social media to share, co-create, and discuss UGC about the topic (Kaplan and Haenlein 2010).

Sharing of case studies on social networks regarding core principles of digital business development, such as the value proposition design relevant in online business models, is taking on increasing significance (Osterwalder and Pigneur 2012). The recognition of entrepreneurial opportunities in the digital space is often based on capitalizing on the best value proposition practices shared on social networks by serial entrepreneurs, who are interested in developing the global digital start-ups ecosystem (Standing and Mattsson 2018). Information related to digital entrepreneurship success stories induces tremendous network effects when shared on social media, as it offers valuable ideas and insights for people looking to embrace an entrepreneurial career (Kraus, Palmer, Kailer, Kallinger and Spitzer 2019).

In-depth analyses of UGC drawn from social media platforms aiming to explore multi-dimensional practices and interactions of stakeholders in digital innovation ecosystems, have received some attention recently (e.g. Suseno, Laurell and Sick 2018). For example, Sorensen and Drennan (2017) suggest that researchers have neglected to examine value-creating strategies in social media-based communities, indicating that only professionals have posted in social networks such valuable practices. Thus, text-mining empirical assessments using thematic analysis of content related to digital entrepreneurship value-creating practices, are needed.

Digital entrepreneurs take full advantage of platform-based open innovation strategies in their business process planning. Social media is extremely useful for open innovation-based entrepreneurial ventures, as it enables consumers to share their knowledge and experiences about how they became co-designers of products and services via online platforms (Hsieh and Wu 2019). Digital technologies and collaborative platforms enable entrepreneurs to design scalable business models which have the capacity to drive rapid growth. Another advantage of

digital technologies in supporting entrepreneurial projects are best-practice sharing processes that encourage experimentation and learning (Zaheer, Breyer and Dumay 2019).

Given the complexity and impact of constantly evolving social media technologies, there is a need for content analyses that are capable of capturing the multifaceted nature of digital entrepreneurship (El Ouirdi, El Ouirdi, Segers, and Henderickx 2015). The value of social media for encouraging digital entrepreneurship is highlighted in extant research focused on thematic analysis of image-based posting (Lowe-Calverley and Grieve 2018), outlining how entrepreneurs' comments in response to a relevant post might provide support to new or aspiring entrepreneurs.

In an attempt to gain valuable insights about social media content focused on the keywords "digital entrepreneurship" and #digitalentrepreneurship (hashtag), a combination of traditional statistical analysis, cluster analysis, text mining, and sentiment analysis techniques can be used to examine the content that makes people engaged when searching relevant posts in this field (He, Tian, Chen, and Chong 2016). Big data capabilities provide behavioural patterns of social media users who engage with online content about digital entrepreneurship, both on the aggregate level, in the form of information about posts, comments and shares related to this topic and on the individual level, in the form of details about the searcher's behaviour over time (Rialti, Zollo, Ferraris and Alon 2019). In the context of big data, text mining techniques that can automatically process social media content related to digital entrepreneurship are required, as they provide opportunities to access and analyse textual open data from social networks (Kayser and Blind 2017).

Generating many concept terms using text analytics algorithms and consequently a deeper look at the social media posts related to digital entrepreneurship terms, has been used by researchers in their attempt to analyse the behaviour of entrepreneurs-turned-influencers,

who post strategies and tips to attract and engage fans and followers (He and Chen 2014). Leximancer, a computer-driven content analysis software, working with its own algorithms, is used in this study to analyse the meanings found in UGC about digital entrepreneurship, by extracting key lexical themes and concepts (Leximancer 2019). The main tasks that can be performed using Leximancer include: the identification of the main topics within a big qualitative dataset, highlighting how topics relate to each other and using word frequency statistics to generate data visualisations (Angus, Rintel and Wiles 2013; Krishen and Petrescu 2017).

To better manage extracted knowledge from the big social media data and offer valuable insights, there is a strong need for researchers to build conceptual models based on thematic analysis (He, Wang and Akula 2017). However, there are also challenges in using social media data in qualitative research. These challenges are mainly related to the big volume of data, the nature of digital texts, the profiles of behaviour on social media sites, and the authenticity of the data, especially when it is retrieved from social networks (McKenna, Myers and Newman 2017). These challenges are acknowledged by this research. Despite these challenges, social media data can proffer for authentic insights as social media UGC is seen as trustworthy by social media users. UGC highlights the manner through which social media users express themselves and communicate with other people in online social media (Smith, Fischer and Yongjian 2012). User reviews focused on digital entrepreneurship on a specific post on Facebook or user videos discussing the features of a digital business model on YouTube, are examples of UGC (Roma and Aloini 2019).

In the context of Social Exchange Theory (Emerson 1976; Liu, Min, Zhai and Smyth 2016), textual data shared on social networks can be categorized into facts and opinions. The exchange represents the interplay between facts, which are objective expressions, and opinions,

which are subjective expressions that describe people's sentiments, appraisals or feelings (Liu 2010). Considered as one of the major research fields in natural language processing (Jeong, Yoon and Lee 2019), sentiment analysis focuses primarily on opinion expressions that approach people's polarity of sentiment. Sentiment analysis in big data of UGC has begun to make use of machine learning, where the classification of sentiment can be formulated as 'deep learning' with three classes: positive, negative, and neutral (Stieglitz, Dang-Xuan, Bruns, and Neuberger 2014).

Methodology

Content, or thematic, analysis is a sense-making enquiry which aims to identify, analyse and report textual consistencies, patterns, clusters and meanings within qualitative data (Attard and Coulson 2012). Such analysis enables the identification of meaning embedded in qualitative material (Angus, Rintel and Wiles 2013). Recent studies which explored big, qualitative data, have used computer assisted qualitative data analysis software (CAQDAS), such as Leximancer, to minimise researcher's bias often present in techniques involving manual 'handling' of data (e.g.: such as coding in QSR NVivo and manual coding) (e.g.: Wilk, Soutar and Harrigan 2019). Researchers have recognised that CAQDAS, such as Leximancer, enable a more objective and efficient examination of the data, driven by the program (rather than the researcher) and with minimal researcher's epistemological influence which might cloud the analysis (Wilk, Harrigan and Soutar 2018).

Wilk, Soutar and Harrigan (2019) compare Leximancer to QSR NVivo and note that when exploring a new construct (online brand advocacy) using both programs, Leximancer identified thematic clusters and patterns within the data that the researchers were not able to identify with QSR NVivo. Consequently, guided by studies such as Wilk, Soutar and Harrigan (2019) and Attard and Coulson (2012), this present study also adopts an inductive approach,

so that the Leximancer algorithm drives the emergent key words, themes (clusters of words) and patterns, rather than categories predetermined by the researcher, such as those required in QSR NVivo or manual text analysis (Wilk, Soutar and Harrigan 2019).

Data collection

Data for this study were collected through the Salesforce Social Studio program, which is one of the key social listening and monitoring platforms used by the industry (Salesforce 2019). The collected data included all globally, publicly available social media UGC featuring the #digitalentrepreneurship (hashtag) and the keywords 'digital entrepreneurship'. A total of 31,017 posts were retrieved over a three month period from 9 August 2019 to 5 November 2019, and included the textual content of the post, the number of comments a post received, the number of followers the post's author has, the following count (the number of contacts the author follows)and the number of post's views. This length of time (3 months) was deemed sufficient for this study, as no further or new insights into the topic were generated due to 'thematic and data saturation' (Gaskell 2000).

As the dataset included 31,017 social media posts, it was deemed to meet the criteria for big data analysis. Big data is said to be characterised by its sheer volume (the amount of data), velocity (the rate at which the data is created), variety (the form of data) and veracity (the validity of the data) (White 2016). Data in our dataset was plentiful (large volume), in various forms (posts with and without virtual cues such as emojis, emoticons, photos, links and videos), and created with velocity (multiple posts within the same time/day for example). Consequently, quantitative data analysis in SPSS was followed by qualitative analysis in Leximancer, a computer assisted qualitative data analysis software (CAQDAS), which enabled an automated examination of the textual content of the big data collected (Leximancer 2019).

Sample overview

Of the total posts in the sample, the majority at 53.49% (n = 16,593), conveyed positive sentiment about digital entrepreneurship, 3.28% (n = 1,017) expressed negative sentiment, and 43.23% (n = 13,407) were neutral. 0.87% of posts were viewed over 100 times (n=269), and 0.08% of posts received over 100 comments (n=26). Further, 3.36% of posts received 1 or more comment (n=1,042) and 2.75% were viewed 1 or more times (n=853). Out of the posts that did receive a minimum of one comment, the average number of comments was 15.80 (*SD* = 90.356), suggesting that 1,042 posts about digital entrepreneurship from the sample received between 1 and 1916 comments. Out of the posts that did receive a minimum of one view, the average number of views was 1,697.69 (*SD* = 13,306.746), suggesting that 853 views about digital entrepreneurship from the sample received between 1 and 240,524 views (Table 1).

Table 1: Descriptive statistics

Variable	N	Min	Max	Mean	SD
Posts' comments	1,042	1	1,916	15.80	90.356
Posts' views	853	1	240,524	1,697.69	13,306.746

Data analysis and results

Qualitative data analysis

The automated, computer-driven, qualitative data analysis with Leximancer is underpinned by an iterative process of seeding word definitions from frequencies and co-occurrences of words within blocks of text (Wilk, Soutar and Harrigan 2019; Angus, Rintel and Wiles 2013; Sotiriadou, Brouwers and Le 2014). In Leximancer, keywords are called

concepts and these are grouped into themes, whereby the themes are based on contextual similarity and are visually represented on the Concept Map (Cretchley, Rooney and Gallois 2010). The themes are colour-heat-mapped from hottest to coolest, where the colour red represents the 'hottest' or most prominent theme and the colour purple is the 'coolest' or least connected theme within the Concept Map. An Insight Dashboard Report provides Prominence Scores 'PS' for concepts and compound concepts, where PS of 1 or more is considered sufficient to identify unique and important characteristics and, for compound concepts, a PS of 3 or more can be satisfactory (Wilk, Soutar and Harrigan 2019).

In this study, tags were seeded and mapped on the Concept Map to better understand discussions around digital entrepreneurship, specifically in relation to the number of followers, following, comments and views attributed to social media posts about the topic of digital entrepreneurship. Most Leximancer settings were left at default, however, some settings were adjusted to suit social media data analysis. For example, the default 2-sentence block setting was increased to 4 sentences. This is because social media communication is fragmented and uses short sentences. Further, the Edit Stop Word list was updated to include evidence words such as 'good', 'great' and 'never', as these words are usually automatically removed from analysis by the Leximancer program. Many academic studies have used this program in a similar fashion and these studies were consulted during this phase of the research (e.g. Wilk, Soutar and Harrigan 2019; Angus, Rintel and Wiles 2013; Sotiriadou, Brouwers and Le 2014).

Leximancer results

Two Leximancer-driven analyses were performed on the big dataset of social media posts discussing digital entrepreneurship.

<u>Social media activity sentiment, followers, views and comment count</u>

The first analysis included all of the social media posts and Leximancer-tags were seeded for 'positive sentiment', 'negative sentiment' and 'neutral sentiment'; and for the number of followers for each post author and for the number of comments for each post. These tags enabled the exploration of: 1. social media activity sentiment towards digital entrepreneurship (sentiment valence), 2. how much engagement there is through comments (comment count) and views (view count) of social media posts relating to digital entrepreneurship; and 3. the impact (if any) of the number of followers each post's author has (follower count) on the social media activity about digital entrepreneurship. This analysis resulted in the Concept Map in Figure 1.
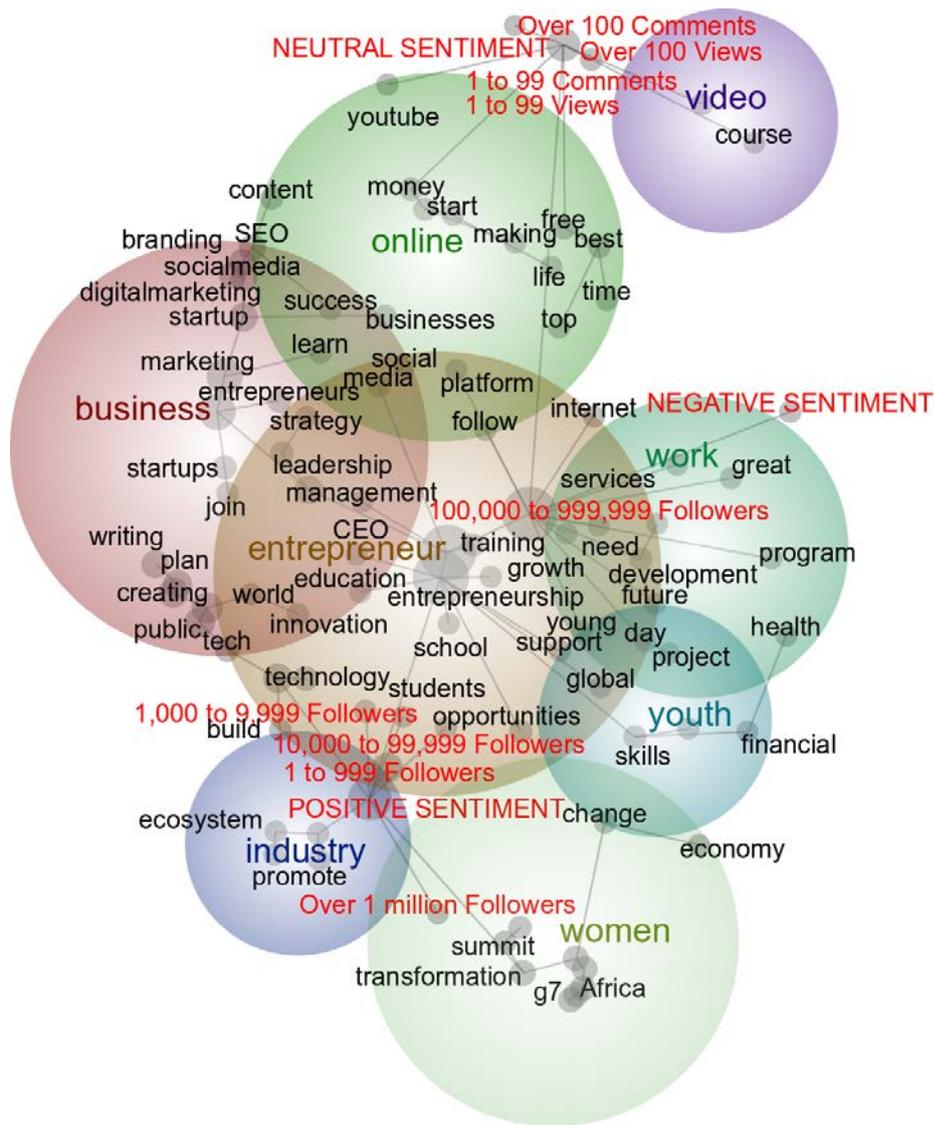
Figure 1: Social media sentiment, the number of comments and the number of followers for all posts about digital entrepreneurship

The data revealed that the most prominent ('hottest', red coloured) theme across all social media posts about digital entrepreneurship was 'business', closely followed by and related to the 'entrepreneur' theme, indicating that the majority of social media posts about digital entrepreneurship were in the context of 'business' and the 'entrepreneur'. The

discussion around business and entrepreneurship was not aligned to a sentiment tag, indicating that all three sentiment valences (positive, negative and neutral) were present and no one sentiment valence took precedence in the posts. When posting about digital 'entrepreneurship' in the context of 'business', other topics included 'branding', 'digital marketing', 'social media marketing', 'content', 'SEO (search engine optimisation)', 'plan', 'leadership' and 'management' which all appeared to be most relevant to digital entrepreneurship in the digital age.

When compared to the Google Trends data over the same period, the words that were most likely to appear in a search with 'digital entrepreneurship' appeared in the following order, with 1 being the most commonly appearing term.

1. Digital media

2. Management

3. Business

4. Course

5. Digital marketing

6. Media

7. Marketing

8. Digital economy

9. Digital transformation

10. Economy

11. India

12. Innovation

13. Master of Business Administration

14. Master's Degree

15. Meaning

16. Social entrepreneurship

The terms in this list are similar to the topics identified in the data based on the social media UGC analysis using Leximancer.

The quotes below illustrate the myriad of topics featured in UGC about digital entrepreneurship; ranging from the development of a digitally-based economy and economic development at a national level, a tool for business growth and strategy; to an individual level, a means of making money and a skill set required for business marketing and growth. There were also posts promoting international events and funding programs aimed at educating people and supporting digital entrepreneurship. For example:

(Topic: Economic development) *"The Foundational Pillars of the Digital Economy 1. Digital infrastructure 2. Digital skills 3. Digital Platforms 4. Digital Financial Services 5. Digital Entrepreneurship."*

(Topic: Business Development) *"Branding it better! Contact us Now! Grow Your Business with us. #marketingstrategy #startup #googlepartner #digital #entrepreneurship."*

(Topic: Individual Skills) *"Builderall Master with Digital Marketing Skillset that will help you make the money you want! Be sure to SUBSCRIBE and take a look at the amazing things you will get when you join."*

Further, the theme of 'entrepreneur' overlapped heavily with the 'business' theme, suggesting there are commonalities in discussion about digital entrepreneurship, specifically, that most discussions included 'business' keywords and 'entrepreneur' keywords. The focus on the 'entrepreneur' in digital entrepreneurship discussion, involved aspects that

entrepreneurs need to focus on to ensure success such as 'strategy', 'leadership', 'management', 'social media platforms'. Notably, much social media activity centred around 'education', 'opportunities', 'innovation', 'school' and 'students'. Here, posts included suggestions as well as exemplars of how digital entrepreneurship is taught in 'schools' and how 'students' are being prepared for digital entrepreneurship 'opportunities', where 'education' plays a role in building 'skills'. Many posts were by education providers showcasing their digital entrepreneurship teaching and learning practices. For example:

"Digital Women Awards celebrate talent, entrepreneurship, innovation and creation among women on the internet. Apply now if you are a woman entrepreneur in the digital space" and

"Embark with us! MASTERS applications OPEN #innovation #aeronautics #aerospace: innovation, project management, entrepreneurship, systems, manufacturing, digital." and

"The youth need to be courageous enough to get what they aspire, whether leadership or even entrepreneurship."

Negative sentiment existed in discussion which pointed out that 'future' 'development' and 'support' of 'youth' globally is needed for effective digital entrepreneurship with a focus on 'financial' and other 'skills'. Table 1 illustrates the most prominent compound concepts (keywords) which co-appeared together most frequently in social media UGC with negative sentiment about digital entrepreneurship. For example:

"Secure and business friendly environment we enjoy today is in our best interest and recently, Digital Entrepreneurship Office has been approved to serve as a centre for

ICT Development; nurture and expand Technovation as well as a bureau for social mobility in solving youth unemployment through the Digital..." and

"Through the Digital Literacy for Employability and Entrepreneurship we give more importance to skill development through entrepreneurship skills and digital skills to underprivileged youth, especially girls."

Posts which displayed 'neutral sentiment' involved video links, most often on Youtube (or cross-promoted on other social media platforms), and involved education about digital entrepreneurship. 'Video' attracted the most engagement through comments (two tags '1-99 comments' and 'Over 100 comments' being pulled by the 'neutral sentiment' tag) and views (two tags '1 to 99 Views' and 'Over 100 Views'); suggesting interactive UGC, such as videos, is needed to drive the digital entrepreneurship agenda online. Interestingly, this discussion centred around online resources and digital entrepreneurship in light of 'making' 'money' and 'best' 'free' resources. This highlights that more resources are needed to support and to ensure the success of digital entrepreneurship and entrepreneurs. Table 1 illustrates the most prominent compound concepts which co-appeared together most frequently in social media UGC with neutral sentiment about digital entrepreneurship. For example:

"If you look at agriculture as an industry, it's one of the least digitalised industries. In this video founder and chair of AgFood Tech. <Name> explains the role of entrepreneurship in agriculture."

Positive sentiment was shared in UGC across all levels of followers. In particular, UGC authors with 'over 1 million followers' were driving positive sentiment on social media. Positive sentiment surrounded discussion about the entrepreneurship 'technology', 'ecosystem' and 'industry, and one which 'promote' entrepreneurship and 'women' 'transformation' of the digital entrepreneurship globally, specifically with initiatives such as

the 'G7' 'summit' in 'Africa' held in August 2019, driving this change. Table 1 illustrates the most prominent compound concepts which co-appeared together most frequently in social media UGC with positive sentiment about digital entrepreneurship. For example:

"#TechMeeting A startup studio for the machine economy that backs, builds, and scales successful ventures, promotes entrepreneurship and feeds growth in Europe's digital economy #blockchain #chooseparisregion #entrepreneurs." and

"Technology that is transformational and not transactional, and the power of the youth to push India's startup movement were the big themes of discussion at the fifth edition of The Economic Times Startup Awards, the most coveted prize for India's thriving entrepreneurship..." and

"As we celebrate #IDRW2019, meet Diarra, a young mother who uses the #RiceAdvice app to help Malian farmers reap the rewards of their hardwork. Check out how #digital tools create job and entrepreneurship opportunities for #ruralwomen." and

"The session discussed key achievements of the G7 and Africa partnership with a special focus on women entrepreneurship, digital transformation, and the fight against corruption."

Table 2. Comparison of top five ranking Compound Concepts between Positive, Negative and Neutral Sentiment

Positive Sentiment: Compound Concepts	PS	Negative Sentiment: Compound Concepts	PS	Neutral Sentiment: Compound Concepts	PS
'transformation' and 'G7'	17.0	'development' and 'skills'	59.0	'video' and 'making'	18.0
'tech' and 'industry'	17.0	'skills' and 'training'	58.8	'money' and 'online'	17.4
'G7' and 'women'	16.2	'time' and 'branding'	50.3	'start' and 'support'	16.1
'women' and 'Africa'	13.5	'money' and 'financial'	44.8	'youtube' and 'learn'	14.3
'summit' and 'change'	10.6	'time' and 'startups'	32.7	'course' and 'free'	13.9

<u>Social media activity analysis of comment count, view count, the number of followers and following count</u>

The second analysis included *all* of the social media UGC and Leximancer-tags were seeded for: the number of followers for each post's author, the following number for each post's author, and the comment and view counts. These tags enabled the exploration of the interplay between social media activity engagement (the number of comments and views) and the number of followers each post's author has (follower count) and the following for each post's author (following count). This analysis resulted in the Concept Map in Figure 2.

Figure 2: Social media activity analysis, comments, views, and the number of followers and following count for
all posts about digital entrepreneurship

Even with the removal of the sentiment tags, this analysis confirmed the findings from the first round, namely the key themes evident in the social media UGC about digital entrepreneurship. Moreover, the Concept Map shows that views and comments are most closely aligned with video content within social media UGC, and such UGC receives most

engagement on social media when it comes to communicating about digital entrepreneurship. The followers and following counts revealed that regardless of the level of connectivity and networking on social media (followers count and following count tags in Concept Map 2), the majority of social media posts about #digitalentrepreneurship were about 'business' and 'entrepreneur'; further highlighting these themes as key to digital discussion in this area. For example:

"Digital Entrepreneurship is probably the best bet for young entrepreneurs with little to no capital." and

"Digital marketing allows small businesses to compete with a much smaller advertising budget." and

"Digital Skills and Entrepreneurship Forum is a platform to equip Young Nigerians with the 21st century skills needed for entrepreneurship and jobs in this fourth industrial revolution that is upon us." and

"Your personal brand is your reputation. And your reputation in perpetuity is the foundation of your career. #startups #business #entrepreneur #digitalentrepreneurship #brand #digital marketing."

Conclusions, implications, limitations and further research

This study revealed the diversity of online discourse found in social media UGC about digital entrepreneurship. Specifically, it is shown that digital entrepreneurship can be viewed from a national and an individual perspective. Digital entrepreneurship is a skill to be taught and learned, a business development and growth strategy, and an economic development tool

for nations. Furthermore, social media UGC is an insightful and information-rich resource for aspiring entrepreneurs, whom through their interactions within social networks, are able to interact with experienced and knowledgeable entrepreneurs, who may act as informal mentors and supporters.

The digitisation of business processes, education delivery, industries and economies, have all provided opportunities for digital entrepreneurs, while highlighting the need for many to "catch up" with the current wave of digital transformation. The #digitalentrepreneurshiphashtag was found to be used in relation to economic development on a national level, in emerging economies such as Africa and India. From the European perspective, digital entrepreneurship was seen as a business growth area, ranging from start-ups to small and medium enterprises. Innovation, technology and start-ups where all terms that were associated with digital entrepreneurship, online companies, ecommerce, fintech and software that form the new digital economy.

Digital entrepreneurship was also linked to marketing activities and channels, encompassing digital marketing, social media and the branding of businesses. There was a strong theme around education and training through online courses being offered as a means of gaining the skills and knowledge required to be a digital entrepreneur and accessing the digital economy. There was also the promise of making money or saving money with free courses.

UGC was plentiful in promoting events and funding programs that assisted in the education of people in their digital entrepreneurship ventures. Topics around youth and women entrepreneurs highlighted the focus on diversity within digital entrepreneurship. When compared to the Google Trends data over the same period, the frequency of words most likely to appear in a search with digital entrepreneurship (hashtag or keyword) included: digital

media, business, course, digital marketing, digital economy, digital transformation, innovation and social entrepreneurship. Not surprisingly the Google Trends insightswere very similar to those obtained through the Leximancer analysis.

Notably, little research has been conducted on how UGC about digital entrepreneurship is influenced by a specific type of context, namely, different types of social media. Except for some notable works (e.g.: Jagongo and Kinyua 2013; Goh, Heng and Lin 2013), prior studies have mostly focused on the analysis of major features of entrepreneurship evident in UGC and the underlying social media users' behaviour, while searching relevant content associated with this topic. However, given the increasing impact of digital entrepreneurship, research into whether and how social media users frame or shape their UGC about digital entrepreneurship across different social media networks, can provide nascent entrepreneurs useful insights on the most suitable strategies to adopt on different social media platforms in order to better engage their clients in brand co-creation processes.

This study demonstrates that UGC about digital entrepreneurship represents a meaningful way of contributing to generating potential entrepreneurs' interest in developing ventures based on ideas and topics shared within online communities. The content analysis underlining the value of comments and discussions around #digitalentrepreneurship can be used as a framework to improve our understanding of how content reflecting these keywords can be turned in sense-giving, creating entrepreneurial opportunities in online niches.

Our study presents valuable insights for understanding the online narrative and sentiment generated via UGC in social media activity around digital entrepreneurship. In this paper we have provided the most recent insights into online discussions which used the #digitalentrepreneurship (hashtag) and/or the keywords 'digital entrepreneurship'. However, the analysis we have provided is based on a snapshot-in-time, visual-first approach, and

presents an opportunity for future research to evolve this work and to present further insights into the state of #digitalentrepreneurship evident in online discussions.

Firstly, the data we have acquired is for the most recent time period, namely from 9 August 2019 to 5 November 2019. We encourage future research to follow our study with a longitudinal, perhaps a 5 to 10 year approach. Such an approach would advance our understanding of the changes and trends occurring in digital entrepreneurship. It would also allow for an exploration of the evolution of the #digitalentrepreneurship (hashtag) use, specifically, how it is used on social media and whether it (the hashtag) or the keywords 'digital entrepreneurship' are used most often in online conversations on this topic; that is, to better understand the linking 'device' for online conversations about digital entrepreneurship.

Secondly, a longitudinal study would inform academia and the industry on how the themes evident in social media activity around digital entrepreneurship, have changed; that is, whether the topics which social media users have been posting about when referring to digital entrepreneurship, have changed over a longer period of time (i.e.: several years rather than months, as is the case in our study). This would be particularly important to better understand the digital entrepreneurship challenges and issues which are talked about by social media users; and whether there are any differences in the challenges and issues faced 5-10 years ago versus currently. Similarly, such a longitudinal approach would enable the identification of success stories and the comparison and contrast of such stories over the 5-10 year period. A longitudinal analysis would enable the assessment of the evolution of social media discussion around digital entrepreneurship and the identification of how online narrative about digital entrepreneurship has changed over the years.

Thirdly, our visual-first analysis, is based on the automated, lexical, thematic analysis of a big data set comprised of social media posts. Future research is encouraged to undertake a

quantitative approach, perhaps a survey of social media users who have actively contributed to or generated social media discussion around digital entrepreneurship. Such research would contribute to better understanding social media users' motivations and attitudes towards digital entrepreneurship and towards engaging in online conversations about digital entrepreneurship. A quantitative analysis would further validate our qualitative findings and provide insights into conceptual relationships of interest around social media users' motivations, perceptions and attitudes towards digital entrepreneurship and towards engaging in social media activity about digital entrepreneurship.

Finally, our study used one computer assisted qualitative data analysis software (CAQDAS), Leximancer, which enabled an automated analysis of the big data set with minimal researcher input. Alike a study by Wilk, Soutar and Harrigan (2018), future studies could perform a similar analysis and use a different CAQDAS, perhaps, QSR NVivo, which requires more involvement from the researcher and more control over the coding of data, to see how the results compare when using different approaches. Such a study would be useful to identify whether any new or other themes emerge in social media activity on this topic.

References

Angus, D., Rintel, S., & Wiles, J. (2013). Making sense of big text: a visual-first approach for analysing text data using Leximancer and Discursis. *International Journal of Social Research Methodology*, *16*(3), 261-267. https://doi.org/10.1080/13645579.2013.774186

Attard, A., & Coulson, N. S. (2012). A thematic analysis of patient communication in Parkinson's disease online support group discussion forums. *Computers in Human Behavior*, *28*(2), 500-506. https://doi.org/10.1016/j.chb.2011.10.022

Barger, V., Peltier, J. W., & Schultz, D. E. (2016). Social media and consumer engagement: a review and research agenda. Journal of Research in Interactive Marketing, *10*(4), 268-287. https://doi.org/10.1108/JRIM-06-2016-0065

Blau, P. M. (1968). Social exchange. *International encyclopedia of the social sciences*, *7*, 452-457. [DOI not available].

Cretchley, J., Rooney, D., & Gallois, C. (2010). Mapping a 40-year history with Leximancer: Themes and concepts in the Journal of Cross-Cultural Psychology. *Journal of Cross-Cultural Psychology*, *41*(3), 318-328. https://doi.org/10.1177/0022022110366105

El Ouirdi, A., El Ouirdi, M., Segers, J., & Henderickx, E. (2015). Employees' use of social media technologies: a methodological and thematic review. *Behaviour & Information Technology*, *34*(5), 454-464. https://doi.org/10.1080/0144929X.2015.1004647

Emerson, R. M. (1976). Social exchange theory. *Annual Review of Sociology*, *2*(1), 335-362.
https://www.annualreviews.org/doi/abs/10.1146/annurev.so.02.080176.002003?journalCode=soc

European Commission. (2015). European Commission (EC), Digital Transformation of European Industry and Enterprises; A report of the Strategic Policy Forum on Digital Entrepreneurship, available from:

http://ec.europa.eu/DocsRoom/documents/9462/attachments/1/translations/en/renditions/native

Gaskell, G. (2000). *Individual and group interviewing*, in M Bauer and G Gaskell, (eds), Qualitative researching with text, image and sound: A practical handbook, pp. 38-56. Sage Publishing Ltd., London, UK.

Goh, K. Y., Heng, C. S., & Lin, Z. (2013). Social media brand community and consumer behavior: Quantifying the relative impact of user-and marketer-generated content. *Information Systems Research*, *24*(1), 88-107. https://doi.org/10.1287/isre.1120.0469

He, W., & Chen, Y. (2014). Using blog mining as an analytical method to study the use of social media by small businesses. *Journal of Information Technology Case and Application Research*, *16*(2), 91-104. https://doi.org/10.1080/15228053.2014.943092

He, W., Tian, X., Chen, Y., & Chong, D. (2016). Actionable social media competitive analytics for understanding customer experiences. *Journal of Computer Information Systems*, *56*(2), 145-155. https://doi.org/10.1080/08874417.2016.1117377

He, W., Wang, F. K., & Akula, V. (2017). Managing extracted knowledge from big social media data for business decision making. *Journal of Knowledge Management*, *21*(2), 275-294. https://doi.org/10.1108/JKM-07-2015-0296

Hsieh, Y. J., & Wu, Y. J. (2019). Entrepreneurship through the platform strategy in the digital era: Insights and research opportunities. *Computers in Human Behavior*, *95*(6), 315-323. https://doi.org/10.1016/j.chb.2018.03.033

Jagongo, A., & Kinyua, C. (2013). The social media and entrepreneurship growth. *International Journal of Humanities and Social Science*, *3*(10), 213-227. http://business.ku.ac.ke/images/stories/research/social_media.pdf

Jeong, B., Yoon, J., & Lee, J. M. (2019). Social media mining for product planning: A product opportunity mining approach based on topic modelling and sentiment analysis. *International Journal of Information Management*, *48*, 280-290. https://doi.org/10.1016/j.ijinfomgt.2017.09.009

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59-68. https://doi.org/10.1016/j.bushor.2009.09.003

Kayser, V., & Blind, K. (2017). Extending the knowledge base of foresight: The contribution of text mining. *Technological Forecasting and Social Change*, *116*, 208-215. https://doi.org/10.1016/j.techfore.2016.10.017

Kraus, S., Palmer, C., Kailer, N., Kallinger, F. L., & Spitzer, J. (2019). Digital entrepreneurship: a research agenda on new business models for the twenty-first century. *International Journal of Entrepreneurial Behavior & Research*, *25*(2), 353-375. [DOI not available].

Krishen, A. S., & Petrescu, M. (2017). The world of analytics: interdisciplinary, inclusive, insightful, and influential. *Journal of Marketing Analytics*, *5*(1), 1-4. https://link.springer.com/article/10.1057/s41270-017-0016-4

Lai, L. S., & To, W. M. (2015). Content analysis of social media: A grounded theory approach. *Journal of Electronic Commerce Research*, *16*(2), 138-152. https://www.researchgate.net/profile/Wai_Ming_To/publication/276304592_Content_analysi

s_of_social_media_A_grounded_theory_approach/links/5753f39008ae10d933755285/Content-analysis-of-social-media-A-grounded-theory-approach.pdf

Leximancer. (2019). Leximancer User Guide: Release 4.5. Available at: https://doc.leximancer.com/doc/LeximancerManual.pdf. Accessed 2020.

Liu, Z., Min, Q., Zhai, Q., & Smyth, R. (2016). Self-disclosure in Chinese micro-blogging: A social exchange theory perspective. *Information & Management*, *53*(1), 53-63. https://doi.org/10.1016/j.im.2015.08.006

Lowe-Calverley, E., & Grieve, R. (2018). Thumbs up: A thematic analysis of image-based posting and liking behaviour on social media. *Telematics and Informatics*, *35*(7), 1900-1913. https://doi.org/10.1016/j.tele.2018.06.003

McKenna, B., Myers, M. D., & Newman, M. (2017). Social media in qualitative research: Challenges and recommendations. *Information and Organization*, *27*(2), 87-99. https://doi.org/10.1016/j.infoandorg.2017.03.001

Nambisan, S. (2017). Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship. *Entrepreneurship Theory and Practice*, *41*(6), 1029-1055. https://doi.org/10.1111/etap.12254

Nzembayie, K. F., Buckley, A. P., & Cooney, T. (2019). Researching Pure Digital Entrepreneurship–A Multimethod Insider Action Research approach. *Journal of Business Venturing Insights*, *11*, e00103, https://doi.org/10.1016/j.jbvi.2018.e00103

Olanrewaju, A. S. T., Hossain, M. A., Whiteside, N., & Mercieca, P. (2020). Social media and entrepreneurship research: A literature review. *International Journal of Information Management*, *50*(2), 90-110. https://doi.org/10.1016/j.ijinfomgt.2019.05.011

Osterwalder, A., & Pigneur, Y. (2012). Designing business models and similar strategic objects: the contribution of IS. *Journal of the Association for Information Systems*, *14*(5), 237-244. DOI: 10.17705/1jais.00333

Rialti, R., Zollo, L., Ferraris, A., & Alon, I. (2019). Big data analytics capabilities and performance: Evidence from a moderated multi-mediation model. *Technological Forecasting and Social Change*, *149,* 119781, https://doi.org/10.1016/j.techfore.2019.119781

Richter, C., Kraus, S., Brem, A., Durst, S., & Giselbrecht, C. (2017). Digital entrepreneurship: Innovative business models for the sharing economy. *Creativity and Innovation Management*, *26*(3), 300-310. https://doi.org/10.1111/caim.12227

Rogers, D. L. (2016). *The digital transformation playbook: Rethink your business for the digital age*. Columbia University Press.

Roma, P., & Aloini, D. (2019). How does brand-related user-generated content differ across social media? Evidence reloaded. *Journal of Business Research*, *96*(3), 322-339. https://doi.org/10.1016/j.jbusres.2018.11.055

Sorensen, A., & Drennan, J. (2017). Understanding value-creating practices in social media-based brand communities. *The Service Industries Journal*, *37*(15-16), 986-1007. https://doi.org/10.1080/02642069.2017.1373098

Salesforce. (2019). Salesforce.com, Accessed on 30 July 2019.

Smith, A. N., Fischer, E., & Yongjian, C. (2012). How does brand-related user-generated content differ across YouTube, Facebook, and Twitter?. *Journal of Interactive Marketing*, *26*(2), 102-113. https://doi.org/10.1016/j.intmar.2012.01.002.

Sotiriadou, P., Brouwers, J., & Le, T. A. (2014). Choosing a qualitative data analysis tool: A comparison of NVivo and Leximancer. *Annals of Leisure Research*, *17*(2), 218-234. https://doi.org/10.1080/11745398.2014.902292.

Standing, C., & Mattsson, J. (2018). "Fake it until you make it": business model conceptualization in digital entrepreneurship. *Journal of Strategic Marketing*, *26*(5), 385-399. https://doi.org/10.1080/0965254X.2016.1240218.

Stieglitz, S., Dang-Xuan, L., Bruns, A., & Neuberger, C. (2014). Social media analytics. *Business & Information Systems Engineering*, *6*(2), 89-96. https://link.springer.com/article/10.1007/s11576-014-0407-5 .

Suseno, Y., Laurell, C., & Sick, N. (2018). Assessing value creation in digital innovation ecosystems: A Social Media Analytics approach. *The Journal of Strategic Information Systems*, *27*(4), 335-349. https://doi.org/10.1016/j.jsis.2018.09.004.

White, D. (2016). Big data: What is it? *Chemical Engineering Progress, 112*(3). https://www.aplu.org/members/councils/student-affairs/csa-presentations/csa-presentations-SF/2014_CampusLabs_BigData.pdf

Wilk, V., Harrigan, P., & Soutar, G. N. (2018). Navigating Online Brand Advocacy (OBA): An Exploratory Analysis. *Journal of Marketing Theory and Practice, 26(*1-2), 99-116. https://doi.org/10.1080/10696679.2017.1389246

Wilk, V., Soutar, G. N., & Harrigan, P. (2019). Tackling social media data analysis: comparing and contrasting QSR NVivo and Leximancer. *Qualitative Market Research: An International Journal*, *22*(2), 94-113. https://doi/10.1108/QMR-01-2017-0021

Zaheer, H., Breyer, Y., & Dumay, J. (2019). Digital entrepreneurship: An interdisciplinary structured literature review and research agenda. *Technological Forecasting and Social Change*, *148*, 119735. https://doi.org/10.1016/j.techfore.2019.119735.