Examining Competing Entrepreneurial Concerns in a Social Question and Answer (SQA) Platform

Irawan Nurhas[1,2][a], Henri Pirkkalainen[3][b], Stefan Geisler[1][c] and Jan Pawlowski[1,2][d]

[1]*Institute of Positive Computing, Hochschule Ruhr West University of Applied Sciences, Bottrop, Germany*

[2]*Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland*

[3]*Faculty of Management and Business, Tampere University, Tampere, Finland*

Keywords: Entrepreneurial Concerns, Challenges to Startups, Topic Modeling, Entrepreneurship, Sentiment Analysis.

Abstract: This study aims to determine the competing concerns of people interested in startup development and entrepreneurship by using topic modeling and sentiment analysis on a social question-and-answer (SQA) website. Understanding the underlying concerns of startup entrepreneurs is critical to society and economic growth. Therefore, greater scientific support for entrepreneurship remains necessary, including data mining from virtual social communities. In this study, an SQA platform was used to identify the sentiment of thirty concerns of people interested in startup entrepreneurship. Based on topic modeling and sentiment analysis of 18819 inquiries in various forums on an SQA, we identified additional questions about founder figures, keys to success, and the location of a startup. In addition, we found that general questions were rated more positively, especially when it came to pitching, finding good sources, disruptive innovation, idea generation, and marketing advice. On average, the identified concerns were considered 48.9 percent positive, 41 percent neutral, and 10.1 percent negative. This research establishes a critical foundation for future research and development of digital startups by outlining a variety of different concerns associated with startup development in the digital age.

SCIENCE AND TECHNOLOGY PUBLICATIONS

1 INTRODUCTION

This study provides a comprehensive insight into people's concerns interested in entrepreneurship and startup entrepreneurs (SEs). Entrepreneurship is a powerful economic and social force (Harb & Shang, 2021; Schöning, 2013; van Stel et al., 2005). As a result, several initiatives have been launched to encourage SEs to establish their enterprises or business models (Ratinho et al., 2020).

Meanwhile, in the knowledge economy and age of virtual communities, the Social Question-and-Answer (SQA) platform is rapidly moving information-seeking behavior toward a more collaborative and personalized question-and-answer experience based on expertise (Choi et al., 2014). Thus, SQA provides access to business networking and open knowledge- and experience-based business and entrepreneurial solutions to SEs with minimal resources (Shneor & Flåten, 2015).

Although numerous studies on SEs have been conducted (Chandra et al., 2016; Puhakka & Ojala, 2021; Saura et al., 2019), it is important to determine the concerns of the discussed topic on a large scale to capture better the different and more detailed aspects (Saura et al., 2019). Therefore, topic modeling (Chandra et al., 2016; Onan et al., 2016; Rehurek & Sojka, 2010) and sentiment analysis (Gao et al., 2013; Hutto & Gilbert, 2014; Onan et al., 2016) methods were applied to the questions collected from an SQA.

By investigating the substance of the issues in an SQA, we set out to identify which competing concerns are most salient and how these concerns are perceived and evolve over time in the face of digitalization. This study contributes twofold: First,

[a] https://orcid.org/0000-0002-2211-8857
[b] https://orcid.org/0000-0002-5389-7363
[c] https://orcid.org/0000-0002-1976-0013
[d] https://orcid.org/0000-0002-7711-1169

Nurhas, I., Pirkkalainen, H., Geisler, S. and Pawlowski, J.
Examining Competing Entrepreneurial Concerns in a Social Question and Answer (SQA) Platform.
DOI: 10.5220/0010661000003064
In *Proceedings of the 13th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2021) - Volume 3: KMIS*, pages 145-152
ISBN: 978-989-758-533-3; ISSN: 2184-3228
Copyright © 2021 by SCITEPRESS – Science and Technology Publications, Lda. All rights reserved

we identify thirty areas of concern that emphasize the most disputed issues involving positive and negative emotions, thus paving the groundwork for future scholarly investigation. Second, this study expands the body of knowledge based on previous research by emphasizing the use of SQA as a valuable element for researching entrepreneurship and confirming the identified SEs' dilemmas through the lens of SQA, spanning a broader range of SE development topics that are not limited by cultural, country-specific, gender, or age boundaries.

As a reminder, the research presentation should adhere to the following structure. In the next section, we present studies on entrepreneurial challenges for startups and the use of data mining techniques to identify entrepreneurial insights. Subsequently, we describe how we conducted our research using topic modeling and sentiment analysis. We then highlight the competing concerns of SEs in our findings section. Next, we address the study's findings, limitations, and recommendations for future research in the following discussion section-and; finally, we provide an overview of the study report.

2 RELATED WORKS

SEs face a variety of challenges in the early stages or throughout their entrepreneurial journey (Giardino et al., 2015; Nurhas et al., 2020; Saura et al., 2019; Wang et al., 2016), including issues related to product and team development, lean process, business model design, and financing, scaling, partnerships, sales, and product and market alignment (Giardino et al., 2015; Wang et al., 2016). These challenges motivate SEs to freely discuss, share, and communicate attached with emotion in order to highlight the elements of success, concerns, or motivations (Saura et al., 2019) that spread across the Internet, the social media community, including SQA (Chandra et al., 2016; Puhakka & Ojala, 2021; Shneor & Flåten, 2015).

Giardino et al. (2015) highlighted that in addition to the team and product-related challenges, technology uncertainties and acquiring the first customer are prominent challenges for software startups that are also relevant to developing a technology-based startup in the digital age. However, in the study based on the analysis of entrepreneurial posts on social media, these issues were not mentioned as essential concerns for success factors for SEs (Saura et al., 2019).

The potential of identifying the topic of concerns of SEs has already been investigated in a study of shared thoughts on Twitter [1], where topics were categorized into business angels, business plans, methodology, tools, projects, jobs, and founders (Saura et al., 2019). Topic modeling and sentiment analysis are two well-known tools (Gao et al., 2013; Onan et al., 2016) for revealing the categories of concern and emotions behind shared thoughts and have also been used to study entrepreneurship and startup based on a large amount of data collected (Chandra et al., 2016; Harb & Shang, 2021; Puhakka & Ojala, 2021; Saura et al., 2019). The methods have already been used in various fields and SQA (Jiang et al., 2018; Kumar et al., 2018).

Uncovering topics and sentiments from SQA is important to identify motivations, areas of interest, and expectations that will raise community awareness of specific issues (Choi et al., 2014). Furthermore, sentiment analysis has been demonstrated to reveal feelings and perspectives that can influence decision-making on particular concerns (Chen et al., 2018; Choi et al., 2014; Johnson, 1990). Consequently, it is critical for triggering discussion, directing, and promoting the selection of policies, strategies, and approaches to entrepreneurial development (Blanchflower & Oswald, 1998; Gifford, 1992; Nurhas et al., 2020; O'Shea et al., 2017).

3 METHODS

Two commonly accepted strategies for doing SQA research are user-based and content-based research (Choi et al., 2014). This research involved the use of a content-based SQA method that focuses on content patterns rather than user characteristics (Choi et al., 2014). The Selenium Python package was used to extract snapshot data at the beginning of the process (Huber et al., 2011). The data was compiled from the text of questions posted in the SQA Quora[2] forums dedicated to startups and entrepreneurship.

We chose Quora as our study platform because it appears to be a promising source for investigating various social phenomena in other studies (Jiang et al., 2018; Wang et al., 2013). Quora is an online community founded in 2009 and headquartered in California, United States of America. Quora focuses on an online open-knowledge that is based on user-generated content. Quora's primary content is a question posed by a user to which other users respond. Users of Quora are expected to use their real

[1] https://www.twitter.com/

[2] https://www.quora.com/

names, though they can choose to remain anonymous. Quora is ranked #356 globally by AlexaPage Rank (as of June 15, 2021) and is available in many countries. Even though Quora's content is open, the company does not offer an official API for retrieving data from its websites.

In the following instance, data collection was done by taking all the data in Quoran SQA on September 21, 2018, removing duplicate questions (N_initial:24802 questions to N_final: 18819 questions), cleaning the data by removing punctuations and stop words, developing labels for different topics as concerns, and identifying the sentiment of each topic to identify competing concerns. The Latent Dirichlet Allocation (LDA) algorithm of Gensim was used as a Python package for clustering topics because it is considered to outperform other known topic models (Harb & Shang, 2021).

The LDA class of statistical language models is used in generative probability computing, which is a subset of statistical computing. LDA makes Topics using word clusters instead of text clusters to understand data better (specifically, in this study, the SQA's questions). In LDA, each topic is a model of a mixture of words. Each word is represented by a Dirichlet distribution coefficient. The model in each topic has the capability to forecast significantly related topics in a document (Blei et al., 2003). For LDA, a predictive likelihood-based approach was utilized to pick out the most optimal number of topics (Chang et al., 2009). Then, the concern was labeled based on the given keywords and the coefficient of each identified topic from the Gensim LDA model. Three to ten relevant keywords of word clusters were used to designate each topic.

Following the identification of each cluster's pertinent topics, the next step is to ascertain the sentiment for each question posed within each cluster. The Vader algorithm (Hutto & Gilbert, 2014) was used to identify sentiment for each issue, which shows not only sentiments and viewpoints from written language but also emojis or spoken emotions (Hutto & Gilbert, 2014; Onan et al., 2016) that are relevant and frequently contained in a text (Gao et al., 2013). All queries that have the same identified topic were grouped individually. The Vader algorithm determines each question's negative, positive, and neutral trends on the same topic. The percentage of sentiments for each issue was derived by comparing the three tendencies to the total number of questions on a specific topic of concern.

Following that, each topic's popularity ratio was determined and organized in descending order (1 to 30, where 30 is the most popular topic in a given year; and 0 for topics not found in a given year, this may happen, for example, if a topic appears after 2010). The popularity ratio value is the percentage of questions asked about a specific topic in a given year compared to the total number of inquiries from all topics. It is compared to the percentage of each topic to detect dynamic changes in topic popularity.

4 RESULT

Based on the log-likelihood value of various learning decay scenarios (0.5, 0.7, & 0.9). Thirty topics with a consistent value amongst diverse scenarios of concern were taken into account. As shown in Figure 1, thirty topics were chosen as the number that can yield the greatest number of topics for the dataset. The selection is based on the maximum feasible log-likelihood value under specific conditions before the log-likelihood value for a larger number of topics drops significantly.

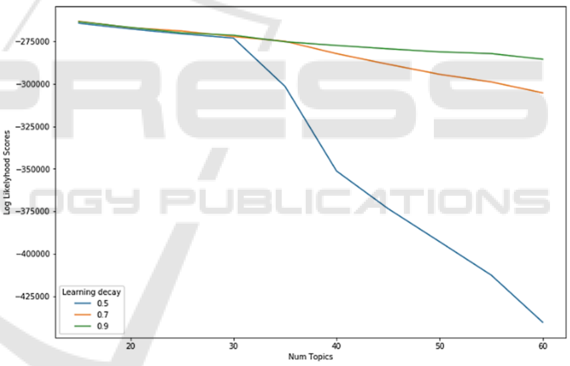

Figure 1: log-likelihood value for a different number of topics.

Table 1 presents the thirty labeled concerns as a result of defining the cluster from LDA. This includes, for example, lean startup, the entrepreneurial process, pitching the idea, scaling the business, and managing company resources, such as founders and co-founders, investments, and location, partnerships, and emerging trends (Giardino et al., 2015; Ratinho et al., 2020; Shneor & Flåten, 2015; Wang et al., 2016).

Overall, sentiment for each concern was positive, with an average percentage of 48.9 percent. All competing concerns for positive and negative were outlined in gray in Table 1. For example, concerns about Pitching and finding good sources with 66% and concern about success keys (64.8%) were

Table 1: List of concerns of SE and the identified sentiment.

Label of concern	Top 3 words and the coefficient	*N*	Percentage of sentiment (%) *(-)*	*(n)*	*(+)*
Lean startup	0.34*"startup"+ 0.18*"lean"+ 0.05*"book"	751	7.1	53.5	39.4
Inspiring Company	0.11*"company"+ 0.12*"grow"+ 0.06*"learn"	743	10.2	42.8	47.0
Founder figure	0.17*"elon_musk"+ 0.05*"time"+0.04*"tesla"	895	24.4	40.2	35.4
Find co-founder	0.18*"startup"+ 0.11*"founder"+0.05*"hire"	756	10.7	44.3	45.0
Success keys	0.19*"entrepreneur" + 0.15*"successful"+ 0.03*"key"	881	10.4	24.7	64.8
User-centric app	0.12*"build"+ 0.08*"app"+0.06*"user"'	658	5.6	43.6	50.8
A year of remote work	0.18*"work"+ 0.07*"year"+0.06*"job"	701	17.1	43.5	39.4
Payment service	0.13*"service" + 0.06*"pay"+0.05*"problem"	672	16.4	32.9	50.7
Pitching	0.12*"innovative"+ 0.09*"idea"+0.06*"sell"	653	6.9	27.1	66.0
Investment capital	0.27*"start"+ 0.11*"business"+ 0.07*"investment"	623	10.1	49.9	40.0
Scaling the business	0.21*"small"+ 0.19*"business"+ 0.06*"scale"	521	8.3	49.7	42.0
Breakeven point	0.25*"make"+ 0.14*"people"+ 0.11*"money"	716	15.6	42.0	42.3
Small business owner	0.22*"business"+ 0.18*"small"+ 0.08*"owner"	540	11.7	45.4	43.0
Disruptive innovation	0.18*"innovation"+ 0.04*"change"+ 0.03*"life"	787	8.8	30.9	60.4
Important items	0.09*"thing"+ 0.07*"important" + 0.06*"open"	647	7.4	41.7	50.9
Find good sources	0.28*"good"+ 0.14*"find"+ 0.07*"software"	447	5.8	28.2	66.0
Funding	0.20*"startup"+ 0.06*"investor"+ 0.05*"invest"	608	9.7	46.4	43.9
Startup location	0.32*"startup"+ 0.32*"tech"+ 0.02*"silicon_valley"	836	8.4	45.2	46.4
Product launch	0.25*"product"+ 0.11*"market"+ 0.08*"launch"	672	8.2	38.5	53.3
Initate new business	0.36*"business"+ 0.08*"run" + 0.08*"start"	513	5.3	35.5	59.3
Risk management	0.27*"strategy"+ 0.04*"fail"+ 0.03*"corporate"	684	13.0	45.2	41.8
Business plan	0.40*"business"+ 0.15*"plan"+ 0.12*"good"	486	5.8	50.6	43.6
Business model	0.12*"business" + 0.10*"model"+ 0.08*"strategic"	591	6.4	50.3	43.3
Marketing advise	0.08*"marketing"+ 0.07*"give"+ 0.07*"website"	531	7.3	32.4	60.3
Validating idea	0.12*"customer"+ 0.08*"development" + 0.05*"idea"	497	10.1	40.2	49.7
Digital business	0.31*"business"+ 0.25*"start"+ 0.10*"online"	431	7.7	45.0	47.3
Emerging markets	0.16*"start"+ 0.13*"business"+ 0.10*"india"	496	9.3	56.7	34.1
Partnership	0.37*"company"+ 0.04*"partner"+ 0.03*" competitive_ advantage"	447	13.9	36.7	49.4
Technology trends	0.11*"technology"+ 0.09*"big"+ 0.08*"industry"	509	9.4	41.3	49.3
Idea invention	0.15*"idea"+ 0.11*"invention"+ 0.10*"good"	527	12.0	24.7	63.4

N: total questions; (-): negative sentiment; (+): positive sentiment; (n): neutral sentiment;

positively perceived. Conversely, some issues are perceived as challenges to SEs development, with concerns about founder figure (24.4%), remote work (17.1%), payment services, and break events (at around 16% each) being perceived as more negative. Based on Table 1 and in light of previous literature on the dimensions of startup challenges, including product, financial, team, and market (Giardino et al., 2015; Ratinho et al., 2020; Shneor & Flåten, 2015; Wang et al., 2016), we identified the following as additional concerns of SEs that need to be considered.

As illustrated in Table 2, international startups and role models are two additional dimensions of

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

challenges that have been overlooked previously and may arise due to Quora's global nature. As a result, people from various countries use the platform to inquire about opportunities, challenges, and success factors associated with developing startups in other locations and require an example of successful development.

The "role models" category is an example of how digital technology can help small and medium-sized enterprises (SMEs) find inspiring stories of role models who are either individuals or organizations to gain insights into how startups in situations similar to their own are evolving.

Table 2: Further concern dimension for SEs.

(Challenge category): definition	Label of concern	Example question from Quora
(international startup): lack of knowledge regarding place developing international startup	Startup location	*"Why has Sub-Saharan Africa failed to produce tech giants like Twitter, Facebook, Apple, Google, Arm (acquired by SoftBank), etc.?"* *"Is Silicon Valley really the best tech startup location?"*
	Emerging market	*"What are the ways to start a tech startup in India while being located in the USA? "*
(Role model): lack of inspiring role model	Founder Figure	*"What Elon Musk Gets Wrong About Leadership?"* *"Why did Steve Jobs push technology forward, even risking his company to fail?"* *"Why are there relatively few female tech startup founders?"*
	Inspiring company	*"I'm 16, and I want to start my own software company like Apple and Microsoft. How and where should I start with?"* *"What do you learn in a small company that you can't learn in a big one?"*

SQA provides a place for direct engagement with other successful SEs when it comes to using SQA to look for role models. They may have a comparable background and can serve as inspiring role models.

As seen in Figure 2, the number of queries for all topics climbed dramatically year after year. Based on the percentage of the ratio value of each year's calculation. We discovered some intriguing patterns of concerns that can help us better grasp how the popularity of a specific issue competes in the SQA.

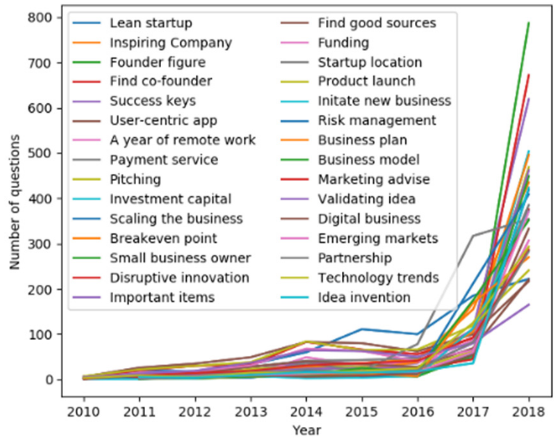

Figure 2: Number of questions per year of topic concerns.

First, we found that some topics were not queried in the first year, 2010. The topics that appeared for the first time in 2011 are pitching an idea, investment capital, scaling business, small business, important resources, finding good sources, startup location, well-prepared business, business model, digital business, emerging markets, partnership, and idea generation. However, some of the concerns receive more attention from SEs to discuss in the SQA.

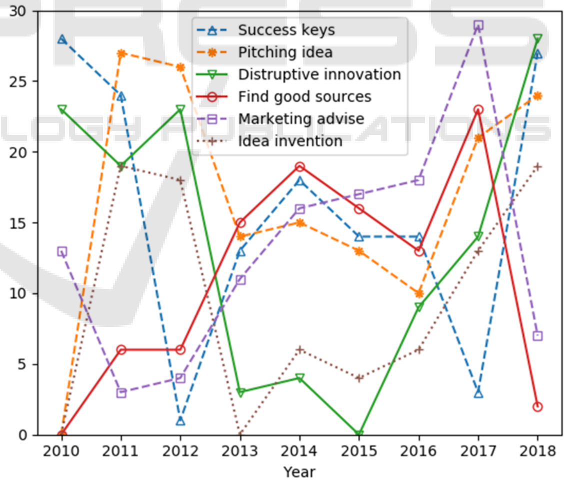

Figure 3: Dynamic changes in popularity over time of positive concerns relative to other concerns.

Next, we present the dynamic changes in popularity of the concerns over time (for Figure 3 to 6, on the x-axis, the years, and the y-axis, the ranking of the concerns relative to each other based on the number of questions asked in that year, the higher the ranking, the more questions asked relative to other concerns in that year). Figure 3 shows the top positive concerns based on Table 1 (pitching, finding good resources, idea generation, marketing advice,

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

disruptive innovation, and keys to success), and Figure 4 shows the top negative concerns and those that have changed significantly over time (Figures 5 and 6).

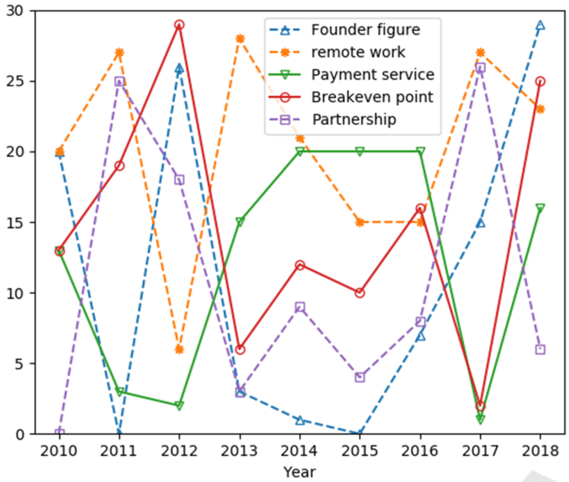

Figure 4: Dynamic changes in popularity over time of negative concerns relative to other concerns.

As shown in Figure 3, the topic of concern about disruptive innovation has received more attention from 2015 to 2018 than other concerns that show more positive sentiment. Meanwhile, the topics of marketing advice and finding good sources fall from the top in 2017.

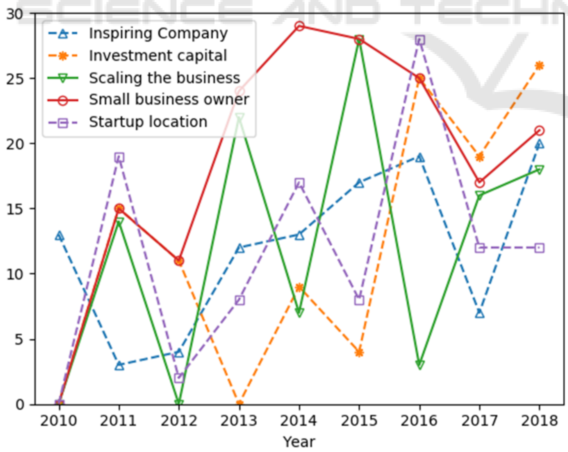

Figure 5: Concerns that show increasing trends in popularity compared to other concerns.

Moreover, we can observe in Figure 3 that there are significant shifts in all concerns in 2014. On the other hand, in Figure 4, in 2015, there were no concerns about the "founder figure," Three years later, the concern was the highest in the group of negative concerns. Thus, overall, the negative concerns group did not show a consistent significant trend over time.

An interesting pattern that shows improvement over time is presented in Figure 5 regarding the concerns of the inspiring company, investment capital, scaling the business, small business owner, and startup location. Those concerns consistently show positive tendencies and can be a signal for providing content learning materials in that regard.

Also, four topics of concern show negative popularity among the graphics: lean startup, user-centered app, product launch, and technology trends. Many reasons might push the concern popularity to be reduced, for instance, the availability of learning sources outside the SQA.

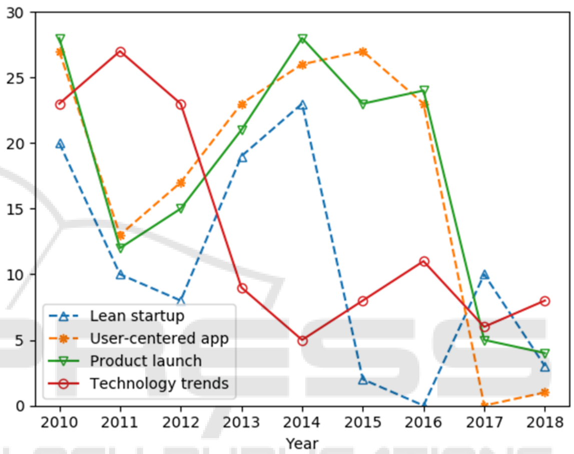

Figure 6: Concerns that show decreasing trends in popularity compared to other concerns.

The following section will discuss the study's implications, including limitations and recommendations for further study.

5 DISCUSSION

Based on the identified concerns from the SQA, further research can use the proposed designation and validate it empirically with SEs based on different types and sizes of businesses, SEs characteristics (Chandra et al., 2016), or the current entrepreneurial journey process. To fit the context of the SEs study, the list of competing concerns may be prioritized, modified, combined, or, if necessary, expanded to include new concerns.

In comparison to earlier research (Giardino et al., 2015; Wang et al., 2016), this study discovered two new dimensions: role model and international startup development. Pitching and product launch were also

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

discussed in the SQA, which shows the importance of startups in the digital age in terms of customer acquisition (Giardino et al., 2015). Technology trends and digital business also shows that the challenges were also prevalent for SEs in the digital age, which was previously discovered only from the social media data of SE status on Twitter (Saura et al., 2019), but from the direct interview and observation study (Giardino et al., 2015; Wang et al., 2016).

Furthermore, the study demonstrates how topic modeling and sentiment analysis may be used to uncover SE concerns based on questions posted on an SQA. We also confirmed SEs' usage of SQA in an open virtual community forum to find preferences and responses to their bewilderment, challenges, and queries as concerns, which had previously been generated by qualitative methods such as surveys, literature studies, and personal interviews with SEs (Giardino et al., 2015; Ratinho et al., 2020; Shneor & Flåten, 2015; Wang et al., 2016).

Nonetheless, this study has a few limitations. First, we only studied one SQA that uses English. Further research can be conducted with other SQA platforms and combine multilingual data collections. Current research on language translation enables topic modeling to mature in English by translating the other language into English before applying the English-based topic modeling. Also, we only looked at the questions and excluded the answers. By supplementing the dataset with responses, additional information can be gathered that will aid in determining the size of the issue at hand.

Furthermore, topics may overlap in meaning or have almost identical meanings; since we only use the log-likelihood value to select the number of topics, other criteria (e.g., coherence value) can also be used. In further studies, the related questions can be included in the analysis to form the label and validate the label with experts and SE directly. Besides, while the study was limited to 2018, based on the identified topics, the LDA model presented in Table 1 with words and the associated coefficient can predict a topic in new documents or questions. Future research may indicate dynamic shifts in the popularity of concern about a particular event, such as before and after the Covid 19 pandemic.

6 CONCLUSION

Our research explores the SEs concern pattern and shows that SQA can be reliable knowledge management and entrepreneurship study tool. We have identified thirty competing concerns that are coupled with SEs' emotional preferences. In the future, it is possible to conduct more empirical explorations based on the thirty issues raised.

ACKNOWLEDGEMENTS

The first author has received a grant from the KPM Committee of the Hochschule Ruhr West to publish the paper with the ID: KPM321024. Additionally, thanks to the Ministry of Culture and Science of the State of North Rhine-Westphalia for the financial support of the Institute of Positive Computing at the Hochschule Ruhr West.

REFERENCES

Blanchflower, D. G., & Oswald, A. J. (1998). What Makes An Entrepreneur? *Journal Of Labor EconomicS*, *16*(1), 26–60.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *The Journal of Machine Learning Research*, *3*, 993–1022.

Chandra, Y., Jiang, L. C., & Wang, C.-J. (2016). Mining social entrepreneurship strategies using topic modeling. *PloS One*, *11*(3), e0151342.

Chang, J., Boyd-Graber, J., Wang, C., Gerrish, S., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. In *Neural information processing systems* (pp. 288–296). Citeseer.

Chen, Y., Yuan, J., You, Q., & Luo, J. (2018). Twitter sentiment analysis via bi-sense emoji embedding and attention-based lstm. In *Proceedings of the 26th acm international conference on multimedia*.

Choi, E., Kitzie, V., & Shah, C. (2014). Investigating motivations and expectations of asking a question in social q&a. *First Monday*.

Gao, W., Li, S., Lee, S. Y. M., Zhou, G., & Huang, C.-R. (2013). Joint learning on sentiment and emotion classification. In *Proceedings of the 22nd acm international conference on information & knowledge management*.

Giardino, C., Bajwa, S. S., Wang, X., & Abrahamsson, P. (2015). Key challenges in early-stage software startups. In *International conference on agile software development* (pp. 52–63). Springer.

Gifford, S. (1992). Allocation of entrepreneurial attention. *Journal of Economic Behavior & Organization*, *19*(3), 265–284.

Harb, Y., & Shang, Y. (2021). Linking information technology and entrepreneurship: A literature review. In *Proceedings of the 54th hawaii international conference on system sciences*.

Huber, M., Mulazzani, M., Leithner, M., Schrittwieser, S., Wondracek, G., & Weippl, E. (2011). Social snapshots: Digital forensics for online social networks. In

Proceedings of the 27th annual computer security applications conference.
Hutto, C., & Gilbert, E. (Eds.) (2014). *Vader: A parsimonious rule-based model for sentiment analysis of social media text.* : Vol. *8*.
Jiang, H., Qiang, M., Zhang, D., Wen, Q., Xia, B., & An, N. (2018). Climate change communication in an online q&a community: A case study of quora. *Sustainability*, *10*(5), 1509.
Johnson, B. R. (1990). Toward a multidimensional model of entrepreneurship: The case of achievement motivation and the entrepreneur. *Entrepreneurship Theory and Practice*, *14*(3), 39–54.
Kumar, A., Praveen, S., Goel, N., & Sanwal, K. (2018). Opinion extraction from quora using user-biased sentiment analysis. In *Information systems design and intelligent applications* (pp. 219–228). Springer.
Nurhas, I., Geisler, S., Ojala, A., & Pawlowski, J. (Eds.) (2020). *Towards a Wellbeing-driven System Design for Intergenerational Collaborative Innovation: A Literature Review*.
O'Shea, D., Buckley, F., & Halbesleben, J. (2017). Self-regulation in entrepreneurs: Integrating action, cognition, motivation, and emotions. *Organizational Psychology Review*, *7*(3), 250–278.
Onan, A., Korukoglu, S., & Bulut, H. (2016). Lda-based topic modelling in text sentiment classification: An empirical analysis. *Int. J. Comput. Linguistics Appl.*, *7*(1), 101–119.
Puhakka, V., & Ojala, A. (2021). A review of international entrepreneurship as part of broader opportunity research: Topic modelling approach. *Empirical International Entrepreneurship*, 89–113.
Ratinho, T., Amezcua, A., Honig, B., & Zeng, Z. (2020). Supporting entrepreneurs: A systematic review of literature and an agenda for research. *Technological Forecasting and Social Change*, *154*, 119956.
Rehurek, R., & Sojka, P. (Eds.) (2010). *Software framework for topic modelling with large corpora*. Citeseer.
Saura, J. R., Palos-Sanchez, P., & Grilo, A. (2019). Detecting indicators for startup business success: Sentiment analysis using text data mining. *Sustainability*, *11*(3), 917.
Schöning, M. (2013). Social entrepreneurs as main drivers of social innovation. In *Social innovation* (pp. 111–118). Springer.
Shneor, R., & Flåten, B.-T. (2015). Opportunities for entrepreneurial development and growth through online communities, collaboration, and value creating and co-creating activities. In *Entrepreneurial challenges in the 21st century* (pp. 178–199). Springer.
van Stel, A., Carree, M., & Thurik, R. (2005). The effect of entrepreneurial activity on national economic growth. *Small Business Economics*, *24*(3), 311–321.
Wang, G., Gill, K., Mohanlal, M., Zheng, H., & Zhao, B. Y. (2013). Wisdom in the social crowd: An analysis of quora. In *Proceedings of the 22nd international conference on world wide web*.
Wang, X., Edison, H., Bajwa, S. S., Giardino, C., & Abrahamsson, P. (2016). Key challenges in software startups across life cycle stages. In *International conference on agile software development* (pp. 169–182). Springer, Cham.