Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 167 (2020) 1880–1888

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Computational Intelligence and Data Science (ICCIDS 2019)

Factors affecting future of work: Insights from Social Media Analytics

Pooja Sarin[a], Arpan Kumar Kar[b], Kiran Kewat[c], P.Vigneswara Ilavarasan[d]

PhD Research Scholar[a],DMS, IIT Delhi, New Delhi-110016, India
Associate Professor[b], DMS, IIT Delhi, New Delhi-110016, India

Abstract

Twitter is one of the most comprehensive sources of public conversations over the world. As a platform, it provides a medium to a large group of audience to express their views and opinions. Owing to the snowballing of the cheaper smartphone market coupled with cheaper data services, the platform has experienced an enormous increase in its use and has become a good medium of influence. Hence, Twitter data i.e. around 1.1 lakh tweets over a period of three months (January to March, 2019) are used to extract the tweets surrounding the discussion of future of work which is the result of the emerging technologies vis-à-vis automation, IoT and Industry 4.0. to understand the perceptions of people, organizations and businesses regarding the same. Furthermore, themes surrounding the discussions are identified and technology enablers that are perceived to bring about a change in the future of work. The social media discussions can strongly influence the opinions and perceptions of people, businesses or organizations regarding Future of Work. The results of sentiment analysis exhibit that there is not much negative sentiment regarding the change in nature of work. Artificial Intelligence and Robotics are identified as the main technology enablers.

Keywords: IoT; automation; future of work; Industry 4.0; technology; digitization; employment; Artificial Intelligence.

1. Introduction

The advent of new technologies always put an impact on the nature of work and this impact is being studied by considering the three viewpoints - the “This time is different” scenario, the “This time is no different” scenario and

* Corresponding author. Tel.+91-8375827170.
E-mail address: sarinpooja.ds@gmail.com

10.1016/j.procs.2020.03.207

the "conservative" scenario [1,6]. All the three scenarios are discussed with respect to employment predictions and changes in corresponding policies along with a literature study of the background history forming the base of research or discussion. While any advancement in technology leads to increased efficiency, productivity, safety and convenience, it also raises questions about its impact on skills, jobs, wages and nature of work itself [21,27,28,30]. For example, automation is expected to ease out work in labor-intensive jobs not only in product-oriented businesses but also in the service sector. Hence, there is always a speculation that automation may lead to loss of such jobs. However, many industry leaders and technology experts believe that this would only open new opportunities – building new job profiles [9]. Hence, the required shift in workforce can be catered to with re-training and re-development of labor [2,3]. In the past, ICT technologies and digitalization have transformed the world of work. Considering the same, it becomes imperative to analyze whether automation, IoT, artificial intelligence and the other emerging technologies also have the capability to do so and how are they perceived to shape the future of work.

The term 'Industry 4.0' circumscribes the fifth Industrial revolution – one that marries advanced manufacturing techniques with the Internet of Things to create manufacturing systems that are not only interconnected, but communicate, analyze, and use information to drive further intelligent action back in the physical world. Several facts have dominated the discussion on the changing nature of work due to the advent of concept of Industry 4.0. Technology is blurring the boundaries of the firm and hence shaping the skills needed for work. The idea of robots replacing workers is striking a nerve [3]. Thus, it has become of prime importance to study and analyze the perception of organizations, businesses and industries with respect to the future of work [5]. This has created a lot of debates surrounding learning, unlearning, relearning, employment, retrenchment and other surrounding issues on employee management in this wave of digital transformation [18,25,28].

Traditionally, any adjustments in the administration approaches were customarily spread and talked about in customary media like papers or news channels, but the recent increase in the use of social media has changed the landscape. These days when individuals need to share anything, social media platforms are the go-to place and Twitter has risen as a stage where individuals are meeting up to talk about their issues, encounters and so on. As of late even MNCs have comprehended the significance of such platforms and continuously attempting to connect with the clients through official Twitter handles. Not only it is a comparatively cheaper way, but at the same time it also allows a quicker spread of information and ensuring it reaches a substantial number of people. As can be seen, boundless surveys and interviews to gauge public suppositions is tormented with various issues like being resource intensive, time consuming, statistic assorted variety, measure of information, validness and so forth. In such a situation, an attempt to utilize Twitter to extract information with respect to their perspectives about the impact of emerging technologies vis-à-vis automation, IoT and Industry 4.0 on the Future of Work are to be comprehended as the main drivers responsible to bring about this change.

1.1 Objectives of the study are as follows

- To identify major topics of discussion surrounding Future of Work (FoW).
- To understand public participation through retweets surrounding FoW tweets.
- To understand the sentiments surrounding social media discussions on FoW.
- To understand the main topics generated among the discussions on FoW.
- To understand the technological enablers of FoW using summarized user generated content.

2. Literature Review

Knowledge extraction through social media has opened new avenues for research in Future of work because of pervasive mobile computing, faster networks and ebullition of social media platforms [4,6,8,10]. Twitter has been significantly used in recent years for research in elections & politics [10], healthcare [8], blockchain adoption & diffusion [11,12], emergency response services [16], consumer psychology & brand management [17] etc.

Understanding of emerging trends and themes on future of work can be understood through current discussions on social media platforms. Artificial Intelligence and its applications have eased the amount of efforts and labor needed to put by humans earlier and has enhanced work dimensions in scalable form. Evolutionary AI algorithms [15] mimicking natural phenomenon has enhanced the computational efficiency as well as opened exciting avenues in automation and Robotics [9].

As there are two sides of every coin, Future of Work [27,28,30] is exhilarating as well as challenging to many. It entirely depends upon how humans are perceiving the technology and their discussions surrounding social media.

Tweets can be a very helpful resource for knowledge extraction regarding Future of Work [10,11,12,13]. Extracted knowledge can be inferred from the themes extracted and analyzed statistically [4].

3. Methodology

About 1.1 lakh tweets are extracted for a period of three months from January 2019 to March 2019. The extracted twitter data is then cleaned to remove the URLs, whitespaces, stop-words, etc. The cleaned data is further used for sentiment analysis i.e. to identify the sentiments of the future of work among the people, organizations and businesses. The polarity of the tweets is also categorized into Positive, Neutral and Negative. Further, the LDA algorithm is used to identify 15 topics with 15 terms each [8,10]. Later, the topics are categorized [19,22,23,20] to identify the themes and different probable scenarios of future of work. Also, the different drivers that are mainly responsible for bringing about the change in future of work are identified and analyzed [18,22,24,29]. Following are the steps undertaken for causal knowledge extraction on future of work using Twitter [7,8,10]:

3.1 Data Collection

Data is extracted from the Twitter API using R (RStudio). Around 1.1 lakh tweets with relevant hashtags and of English language are collected. As IoT, Industry 4.0, Automation and Artificial Intelligence are the emerging technologies which are expected to shape the future of work, corresponding hashtags along with that of Future of Work are used for data extraction. Hashtags which are undertaken for data collection are as follows #futureofwork, #industry4.0, #iot, #automation, #ai

3.2 Data Cleaning

The extracted Twitter data is cleaned to remove the punctuations, numbers, stop-words, white spaces, URLs, etc. so that it could be used for further analysis. The cleaned data was then used for Descriptive Analysis, Sentiment Analysis and Topic Modelling.

3.4 Descriptive Analysis

The cleaned data is used to construct the Word-cloud to understand the most frequent words in the data sample, where the word with the highest frequency will have the highest font size and so on. In addition to this, a Retweet analysis is performed to understand when the frequency of tweets is the highest and the ratio of original tweets to retweets at any given point of time.

The ‘word-cloud’, ‘ggplot2’, ‘readr’, ‘lubridate’, ‘syuzhet’ and ‘plyr’ libraries are used for the Descriptive Analysis.

3.4 Sentiment Analysis

The cleaned data is then used to identify the polarity of the tweets by calculating the sentiment score for every individual tweet. The polarity of the words in a tweet is decided by using a collection of positive words and negative words. Thus, for every tweet the positive and negative scores are calculated which are further used to calculate the Sentiment score as follows:

Sentiment Score = Positive Score – Negative Score.

3.5 Topic Modelling

The same corpus is used to form the TDM (Term Document Matrix). Here, 15 topics with 15 words each are identified using the LDA Algorithm and Correlated Topic Model (CTM). The ‘topic-models’ library is used for this purpose.

3.6 ANOVA to understand Technology Enablers for Future of Work

On analyzing the different topics and their corresponding themes from the extracted tweets using Topic Modelling, the different technologies that may act as or are perceived to act as the enablers of Future of Work are identified. ANOVA is a statistical technique that is used to determine whether the means of two or more samples or groups are significantly different from each other. It is used to check the impact of one or more variables or factors by simply comparing the means of the samples or groups. The F-statistic or F-ratio measures whether the means of different groups are similar or not. Lower the value of F-ratio, more similar are the means of the groups and vice versa.

For checking which of the groups or samples have different means, the Bonferroni approach is used and t-test (Two sample assuming equal variances) for every pair of sample or group is performed. This will help in determining which of the samples or groups belong to the same population.

4. Results & Discussion

Knowledge extraction from twitter data is performed by steps mentioned below:

4.1 Descriptive Analysis

4.1.1 Word-Cloud

Exploratory data analysis is performed by constructing a word-cloud with the sample data of approximately 1.11 lakh tweets. We created the word-cloud of maximum 250 words with a minimum frequency of 100 ranging across all the tweets. The words with the highest frequency have the maximum font size while those with low frequencies have a decreased font size.

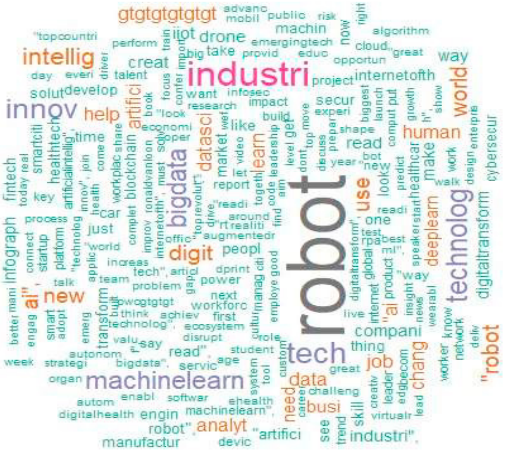

Figure 1: *Future of Work - Word Cloud of all tweets*

From the word-cloud, it is very much evident that the words ‘robot’, ‘machine learning’, ‘industry’, ‘technology’, ‘innovation’, ‘bigdata’, ‘ai’, ‘data science’, etc. are some of the most used words in context to Future of Work. Thus, it can be inferred that as per the perception of people, businesses or organizations, Future of Work will not only be greatly impacted by the emerging technologies but also may depend upon the type of industry and the scope of innovation within a specific industry.

4.1.2 Re-tweet Analysis

Retweet Analysis is performed to understand the retweet ratio, which is the ratio of retweets to total retweets for any given time frame (here for the collected sample data). This retweet ratio can be used to determine the level of social influence with respect to the opinions and views regarding the Future of Work

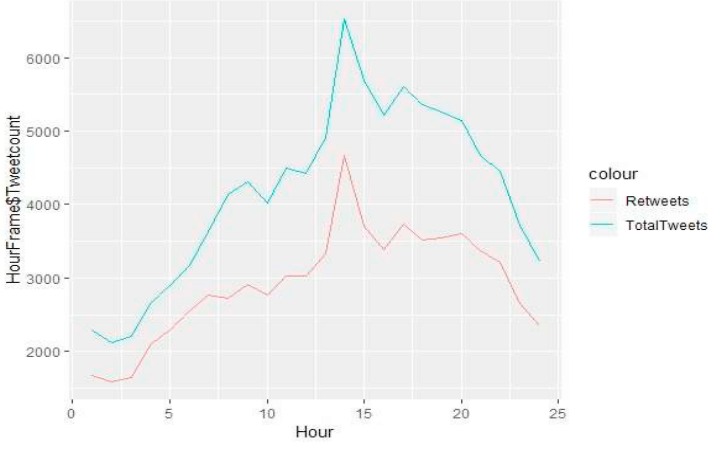

Figure 2: *Retweet Analysis*

From the graph, it can be understood that most of the tweets at any point of time were retweets. Also, the retweet ratio is the highest between the 12:00 to 14:00 hours of any specific day. On analyzing the plot, it is seen the perception of people, businesses or organizations is highly influenced by the content related to Future of Work on social media (Twitter), to which they may agree or disagree. The Retweet Analysis coupled with Sentiment Analysis gives a fair picture of the level of social influence along with the expressed sentiments

4.2 Sentiment Analysis

Sentiment Analysis is used to identify and categorize opinions into positive, negative or neutral. The sentiment score for each tweet is calculated and then aggregated them into positive, negative and neutral categories. Any tweet with a sentiment score > 0 is positive, sentiment score < 0 is negative and sentiment score = 0 is neutral.

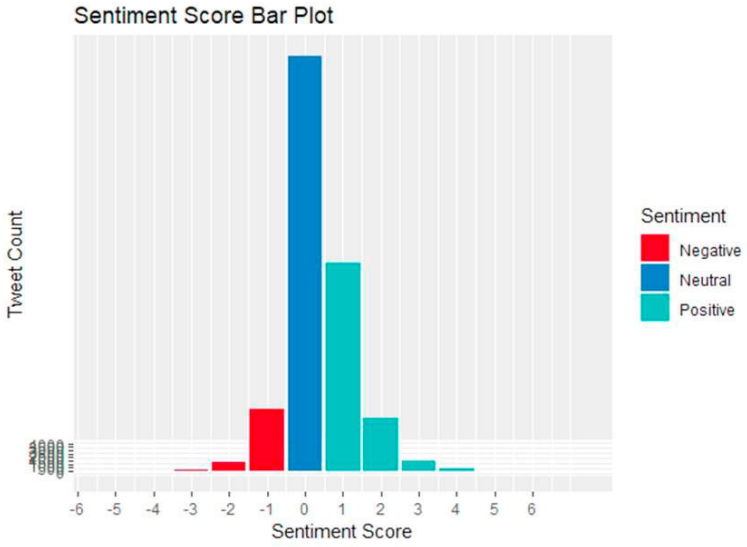

Figure 3: *Sentiment Score Bar Plot*

Thus, majority of the collected tweets were neutral, followed by positive and then negative. From the neutral behavior of the tweets, we can infer that maximum of the tweets may have mixed perceptions about Future of Work. For example, Industry 4.0 in the manufacturing industry may be termed as positive for increasing efficiency but at the same time it may be considered negative for reducing labour. Thus, the final perception becomes neutral.

Also, there are very few tweets that exhibit high negative sentiment scores and even comparatively lower than the number of high positive sentiment tweets. This indicates that the opinions are largely without any biases

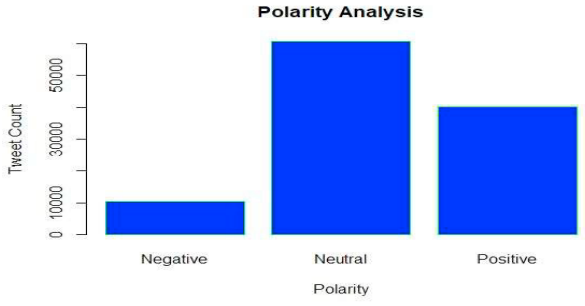

Figure 4: Polarity Analysis – Future of Work

From the above graph, it is seen that out of approximately 1.11 lakh tweets, approximately 60000 tweets exhibit neutral sentiment, followed by approximately 40000 positive sentiments and approximately 10,000 exhibit negative sentiments.

4.3 Topic Modelling

Topic Modelling is used to capture or identify the different themes that may be exhibited in a group of text or documents. The LDA (Latent Dirichlet Algorithm) and CTM (Correlated Topic Model) are used to identify the different themes by grouping the tweets into 30 topics of 15 words each [4,7,8,10]. The list of topics thus obtained are as follows:

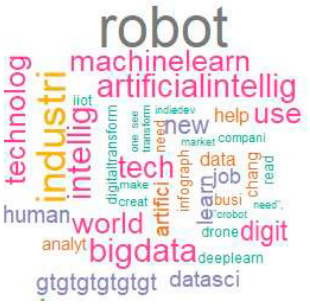

Figure 5: *Future of Work Topic List*

4.4 ANOVA to understand Technology Enablers for Future of Work

On analyzing the different topics and their corresponding themes from the extracted tweets using Topic Modelling, different technologies that may act as or are perceived to act as the enablers of Future of Work are identified. The different technologies are: "Robotics, Artificial Intelligence, Machine Learning, Data Science, Big Data, Digitalization, Automation". All these enablers are rated using the 30 topics and further statistically analyzed them using ANOVA (Analysis of Variance) and t-tests. The result of the One-Way ANOVA analysis can be shown as follows:

Table 1: *One Way ANOVA for Technology Enablers*

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Robotics	30	27	0.9	0.093103
AI	30	30	1	0
Machine Learning	30	21	0.7	0.217241
Data Science	30	16	0.533333	0.257471
Digitalization	30	13	0.433333	0.254023
Automation	30	16	0.533333	0.257471
Big Data	30	20	0.666667	0.229885

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	7.657143	6	1.27619	**6.823529**	1.31E-06	**2.143453**
Within Groups	37.96667	203	0.187028			
Total	45.62381	209				

Here, the value of F-statistic (6.823529) is greater than the F-critical value for the selected level of significance (α=0.05). Also, the p-value is much less than α. Thus, it can be said that at least one of the groups has a different mean and belongs to a different population.

Using the Bonferroni approach, it is tried to find out which of the groups represent different populations. T-tests are performed for every pair of groups and compare their corresponding p-values.

Table 2: *p-values of t-tests (Two sample assuming equal variances) for Technology enablers*

	Robotics	*AI*	*Machine Learning*	*Data Science*	*Digitalization*	*Automation*	*Big Data*
Big Data	0.014170599	0.000170251	**0.39289666**	**0.149925936**	0.035652737	**0.149925936**	
Automation	0.000627815	2.45339E-06	**0.09519483**	**0.5**	**0.223437386**		
Digitalization	2.90506E-05	3.73263E-08	0.018814875	**0.223437386**			
Data Science	0.000627815	2.45339E-06	**0.09519483**				
Machine Learning	0.027022985	0.000416559					
AI	0.038927877						
Robotics							

From the individual t-tests performed for every pair of technology groups, it is seen that the technologies such as Big Data, Data Science, Automation, Machine Learning and Digitalization belong to the same population. On the other hand, technologies such as AI and Robotics belong to an entirely different population. Thus, it can be said that AI and Robotics are perceived to have a significant effect on the Future of Work.

5. Conclusion

The study attempted to understand the factors that are mostly discussed and are of critical concern among stakeholders surrounding FoW. In this context, social media analytics has been applied on user generated content to highlight the following findings across 1.1 lakh tweets.

- The study supports the notion that the social media discussions can strongly influence the opinions and perceptions of people, businesses or organizations regarding Future of Work.
- The results of sentiment analysis exhibit that there is not much negative sentiment regarding the change in nature of work.
- The neutrality of majority of the tweets indicates that the emerging technologies are perceived to have both positive as well as negative impact which may balance out each other in future. However, high positive sentiments otherwise indicate that there is an optimistic approach towards the changing nature of work due to the emerging technologies. The high number of retweets means that the discussions are highly influenced by the opinions and perceptions of major influencers.
- The themes identified from topic modelling identified Artificial Intelligence and Robotics as the main technology enablers. However, many literature studies state the type of industry also plays a major role and is a key facilitator in driving this change.

This exploration paves the path for a much deeper understanding from different stakeholders surrounding drivers and factors which may affect the FoW in the future.

References

[1] De Bruyne, E., & Gerritse, D. (2018). "Exploring the future workplace: results of the futures forum study." Journal of Corporate Real Estate, 20(3), 196-213.

[2] Jeff Schwartz, J. H. (2019, February). "Reframing the Future of Work. Retrieved from MIT Sloan Management Review" https://sloanreview.mit.edu/article/reframing-the-future-of-work/

[3] World Bank. (2019). "World Development Report 2019: The Changing Nature. Washington, DC: World Bank: Creative Commons Attribution CC BY 3.0 IGO."

[4] Khan, Aamera ZH, Mohammad Atique, and V. M. Thakare.: "Combining lexicon-based and learning-based methods for Twitter sentiment analysis. In: International Journal of Electronics, Communication and Soft Computing Science & Engineering" IJECSCSE: 89 (2015)

[5] James Ware, Charles Grantham, (2003) "The future of work: Changing patterns of workforce management and their impact on the workplace", Journal of Facilities Management, 2(2), 142-159.

[6] Ville-Veikko Pulkka, (2019) ""This time may be a little different" – exploring the Finnish view on the future of work", International Journal of Sociology and Social Policy, 39(1/2), 22-37.

[7] Lakhiwal, A., and Kar, A.K. (2016). "Insights from Twitter Analytics: Modelling Social Media Personality Dimensions and Impact of Breakthrough Events. In: Conference on eBusiness, e-Services and e-Society" Springer International Publishing, 533-544.

[8] Grover, P., Kar, A. K., & Davies, G. (2018). "Technology enabled Health"–Insights from twitter analytics with a socio-technical perspective". International Journal of Information Management.

[9] Block, K. (2018, March 05). "New technologies must create jobs, not destroy them. Retrieved from World Economic Forum" https://www.weforum.org/agenda/2018/03/new-technologies-must-create-jobs-not-destroy-them/

[10] Grover, P., Kar, A.K., Dwivedi, Y.K. & Janssen, M. (2018). "Polarization and Acculturation in US Election 2016 outcomes – Can Twitter Analytics predict changes in Voting Preferences". Forthcoming in Technology Forecasting and Social Change.

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

[12] Grover, P., Kar, A.K., Janssen, M (2019). "Diffusion of Blockchain Technology– Insights from Academic Literature and Social Media Analytics". Journal of Enterprise Information Management

[13] Grover, P., Kar, A.K. & Ilavarasan, P.V. (2019)." Impact of Corporate Social Responsibility on Reputation – Insights from Tweets on Sustainable Development Goals by CEOs". International Journal of Information Management.

[14] Grover, P. & Kar, A.K. (2019). "User Engagement for Mobile Payment Service Providers – Introducing the Social Media Engagement model". Journal of Retailing and Consumer Services.

[15] Aswani, R., Kar, A. K., Ilavarasan, P. V., & Krishna, R (2018). "Solving Location Based Inventory Routing Problem in E-Commerce Using Ant Colony Optimization." Lecture Notes in Computer Science, 111.95, 557-566.

[16] Singh N, Roy N, Gangopadhyay A (2018). "Analyzing the Sentiment of Crowd for Improving the Emergency Response Services" IEEE International Conference on Smart Computing

[17] W Read, Robertson, R Mc Quilken, A S Ferdous (2019). "Consumer engagement on Twitter: perceptions of the brand matter" European Journal of Marketing

[18] He, Zhang, Tian, Tao, & Akula, (2019). "Identifying customer knowledge on social media through data" Journal of Enterprise Information Management Vol. 32 No. 1, 2019 pp. 152-169

[19] Ahmed, W. and Lugovic, S. (2019), "Social media analytics: Analysis and Visualisation of news diffusion using NodeXL", Online Information Review, 43(1), 149-160

[20] Urrutikoetxea Arrieta, B., Polo Peña, A. and Martínez Medina, C. (2019), "The moderating effect of blogger social influence and the reader's experience on loyalty toward the blogger", Online Information Review, 43(3), 326-349.

[21] Zhao, Y., Chen, Y., Zhou, R. and Ci, Y. (2019), "Factors influencing customers willingness to participate in virtual brand community's value co-creation", Online Information Review, 43(3), 440-461

[22] Timothy R. Hannigan, Richard F. J. Haans, Keyvan Vakili, Hovig Tchalian, Vern L. Glaser, Milo Shaoqing Wang, Sarah Kaplan and P. Devereaux Jennings (2019), "Topic Modeling in Management Research: Rendering New Theory from Textual Data" Academy of Management Annals 13(2), 586–632

[23] Ann Devitt, Khurshid Ahmad (2013), "Is there a language of sentiment? An analysis of lexical resources for sentiment analysis" Language Resources and Evaluation, 47(2) (Spring 2013), 475-511

[24] Hsinchun Chen, Roger H. L. Chiang and Veda C. Storey (2012), "Business Intelligence and Analytics: From Big Data to Big Impact", MIS Quarterly 36(4), 1165-1188

[25] Sandra Gonzalez-Bailon & Georgios Paltoglou (2015), "Toward Computational Social Science: Big Data in Digital Environments" The Annals of the American Academy of Political and Social Science 659, 95-107

[26] David A. Scheweidel & Wendy W. Moe (2014), "Listening In on Social Media: A Joint Model of Sentiment and Venue Format Choice Journal of Marketing Research" 51(4), 387-402

[27] Sara C. Walton, Paula Marie O'Kane, Diane Rongo Ruwhiu and Virginia Ilene Cathro (2016), "The 'Future of Work': Developing Plausible Scenarios on Work and Business Environments in 2030." Academy of Management Proceedings 2016(1)

[28] Julia Schlegelmilch and Evgenia Lysova (2018), "Digital Nomads and the Future of Work" Academy of Management Proceedings Vol. 1

[29] Paolo Rosso, Marcelo Errecalde, David Pinto (2013), "Analysis of short texts on the Web" Language Resources and Evaluation, 47(1), 123-126

[30] Stefano Scarpetta (2019), "The Future of Work: Advancing Labour Market Resilience" Journal of International Affairs, The Fourth Industrial Revolution,72(1) 51-56