

Lda Topic Modelling On Job Advertisements To Analyze Transformation Of Traditional Quality To Quality 4.0

Junaid Ahmed¹, Dr. Hakeem Ur Rehman², Dr. Sajjad Ahmad Baig³

¹Institute of Quality & Technology Management, University of the Punjab, Lahore, Pakistan.

²Assistant Professor, Institute of Quality & Technology Management, University of the Punjab, Lahore, Pakistan.

³Associate Professor, Faisalabad Business School, National Textile University, Faisalabad, Pakistan.

ABSTRACT

Quality management is also evolving since its inception. However, during a certain interval, it evolves significantly. The introduction of Industry 4.0 again provides the required thrust and quality evolve to quality 4.0. This research aims to analyze the jobs description of quality professionals and identify the key themes to determine that traditional quality (TQM) is transformed into quality 4.0. The research use Indeed job listing against quality 4.0, quality in industry 4.0, digital quality, and smart quality keywords. Orange 3.28.0 is used to apply LDA topic modeling and subsequent analysis. The top ten topics in the job descriptions are Environmental Responsibility, Digital Responsibility, Digital Skills, Digital Knowledge, Data Solution, Quality Training, Quality Knowledge, information excellence, software, and management skills. Information Excellence is in line with the information quality concept. Digital responsibility was not highlighted previously. The job advertisement provides insights about the organization intension and plans and analysis of the job description identified that now the organizations are shifting their intention from traditional quality to quality 4.0 because the no topic identified in this research related to traditional quality like process improvement, process control, etc. therefore, this study concluded that traditional quality is transformed into quality 4.0.

Keywords Quality 4.0; Industry 4.0; Digital Quality; Quality Evolution.

INTRODUCTION

Management approaches are also transformed with new developments that directly or indirectly influence these practices either positively or negatively. Companies' top management visit frequently worksites to get a fresh update but now they see all changes live on their screen with the help of dashboards (Gara et al, 2021). Similarly, meetings are also transformed and now everyone can join a meeting without physically moving to a central location (Rubinger et al, 2020). The management discipline also transformed with time to meet the requirement.

Germany introduce the 4th industrial revolution in 2011 in an international exhibition and provide a foundation for the industry leaders, researchers to adopt and analyze the industry 4.0 (I 4.0) (Kodama, 2018). The I 4.0 is the first industrial revolution that is launched in a planned way. This revolution is based on physical-cyber interaction and vertical and horizontal integration. This revolution transforms the process into an automated process where robotics and sensors control the entire process without human errors and Artificial intelligence analyzes the real-time data (Nica & Stehel, 2021). In case any variation is observed, AI can guide actuators to readjust the process.

This revolution transforms factories into smart factories where computer vision inspects products and AI can manage processes (Villalba-Diez et al, 2019). This new approach eliminates human errors, provides more control over the process, enables the collection of data in real-time, and takes a decision within no time. All these features lead to a zero-defect environment.

However, this transformation also opens a new horizon for quality management because these initiatives provide more precise and real-time data with the ability to analyze and make a decision immediately (Durana et al, 2019). These features are always desired of the good quality system but due to limited technology adoption and organization was one of the key hindrances. Industry 4.0 brings a twist and now organizations are adopting technology for overall benefits and sustainability (Wang et al, 2020).

Now it is important to analyze that organizations are now working on the adoption of quality 4.0 to equip their quality departments. This research identifies the adoption intention based on topic modeling on the job advertisement.

LITERATURE REVIEW

At the start, quality is simply inspection to determine it is good to deliver to the customer or not (Zonnenshain & Kenett, 2020). This approach is fundamentally known as quality control. Later on, the financial element of the quality adopted and poor quality is considered a financial loss. At this stage, quality takes more intension in both practitioners and researchers. This era of quality is also termed quality assurance (Zonnenshain & Kenett, 2020).

Quality is the most prominent word in the industrial revolution of the 20th century and a key determinant for the success and sustainability of the organization. After world worldwar2. Japanese extensively work on quality with the help of American experts and develop a quality culture country-wide. This culture enables Japanese firms to lead the global market after two

decades of effort (Tsutsui, 2019). Although American Experts introduce the concept of scientific quality in Japan, after two-decade, they knock out America from competition based on quality. The more scientific approach adopted in this era and statistical process control are widely adopted to control quality in various organizations.

Quality assurance is more about process and design quality to achieve conformance quality (Feldman, 2005). The third evolution in quality is the development of quality as a holistic discipline. this era of quality is termed total quality management. It is more focused on employee involvement, empowerment, and continuous improvement (Arora & Gupta 2020).

This era started the implementation of the computer in quality and termed computer-aided quality was used (Haynes & Walker (1987). With time, the use of sensors to control processes increased. This adoption equipped quality in comparison to other disciplines and enable the calculation of variance with ease and quickly. However. Still, more of the data is limited to samples and patient experience is too limited considered to estimate customer requirements and satisfaction (Sony et al, 2021).

Quality management in the era of industry 4.0 focused on the discovery of the data and work on insights rather to use traditional root cause analysis (Radziwill, 2018). Organizations can find new insight by augmenting human intelligence with new technologies (Radziwill, 2018). Quality professionals should have the skills to determine what, how, and why data should be used to achieve higher quality (ASQ, 2018).

Radziwill (2018) argue that the following seven tools can play a vital role to manage quality in industry 4.0:

- **Artificial intelligence:** language processing, computer vision, chatbots, navigation, personal assistants, robotics.
- **Big data:** infrastructure (such as Hadoop, MapReduce, NoSQL, and Hive databases), ability to access, manage, and analyze large data on the remote server.
- **Blockchain:** improved transparency and suitability of transactions, monitoring conditions so transactions.
- **Deep learning:** complex pattern recognition, image classification, time series forecasting, creating sound and art, text generation, adjusting images based on heuristics, and creating fictitious video from real video
- **Enabling technologies:** cloud computing, affordable sensors, and actuators, open-source software, mixed reality, AR, VR, 5G networks, data streaming (such as Kafka and Storm), IoT, IPv6.
- **Machine learning:** email spam filters, fraud detection, text analysis, recommendation systems, forecasting, classifying objects into groups
- **Data science:** Power to bring heterogeneous data sets together for making

predictions, performing classifications, finding patterns in large datasets, reducing large sets of observations to most significant predictors, applying sound traditional techniques (such as visualization, inference, and simulation) to generate viable models and solutions. The benefits of the statistical process control increase when the automated analysis is applied as real-time data helps to take spontaneous reactions whenever any deviation is observed in the data. In addition to this, centralization of the data and visibility across the board provides an opportunity to keep aware of all stakeholders at every stage of the process, from start to end. This will help the organizations to increase output by saving time in decision-making.

Srinivasan and Kurey (2014) highlighted the cost of quality data and concluded that companies having high-quality culture save on average \$ 350 Million in comparison to those having poor quality culture. In the recent transition from traditional manufacturing to industry 4.0, there are many challenges for the quality professional to face such as lack of knowledge and how it will influence the existing Quality Management System (Jacob, 2017). Watson et al. (2018) also concluded that the recent revolution brings new challenges and opportunities for quality professionals such as digital diversity, adoption of new technologies, integration of technology into the human sphere.

It is evident from the current literature that quality 4.0 is not a simple addition of one or two factors but an entire shift on a new paradigm with its tools and techniques. Therefore, if organizations are intended to adopt this new quality management system, they hired skilled people to bring change to the organizations. This study identifies the intention to adopt quality 4.0 through analysis of the job advertisements.

RESEARCH METHODOLOGY

Text analytics is widely adopted in recent research (Kim et al, 2017). Text analytics is based on machine learning to analyze the text and determine valuable insights. Topic modelling is most frequently adopted in recent research to identify the key topics in the text. Among all topic modelling approaches, Latent Dirichlet Allocation (LDA) is widely adopted by researchers (Tong & Zhang, 2016; Kolini & Janczewski, 2017; Asmussen, & Møller, 2019; Negara et al, 2019; Jelodar et al, 2019). Therefore, this research also adopted LDA topic modelling to explore the key topics.

Orange 3.28.0 is used in this study to apply LDA topic modelling. Orange is an open sources software freely available. It is based on python but it provides a graphical user interface.

Data is collected from the indeed website. Indeed, is the world's largest job portal (Polner, 2021). Jobs against keyword quality 4.0, Quality in industry 4.0, Digital Quality, and Smart Quality. Total 76 jobs appear on these four keywords. All are included in this study to get more generic insights. From this advertisement, the only detailed job description was extracted and converted into PDF files for analysis.

Data is imported in Orange through important functions. In the next step, the data is converted into a corpus for further processing. Preprocessing is a process that transforms raw data into valuable and comparable data before applying any test. Text analytics techniques are case sensitive, therefore, at first, all text is converted into lowercase. All punctuations and URLs are also removed from the text. In the next step, all text is converted into single-word tokens. Now each file contains token not text. These tokens are further refined by removing general words and structural words from the text. These words are known as stop words. In the first step general stop word list for the English language is used and in the second step customized stop-word list is used to further remove topic-related unnecessary words. After removing words, white spaces are also removed. Now some words appear more or less in every job advertisement that means they are common words and inclusion of those words will not contribute any value. Similarly, some words appear in below 10% of documents that indicate their relevance with the topic is very low. Therefore, the words that appear in more than 90% and less than 10% of jobs advertisement are removed from the corpus. The next step is the conversion of words into base words to increase comparability. The last step is converting tokens into sets of desired numbers. This feature is known as N-gram. In this study all tokens are converted into a set of single and two words. Now the text is converted and refined for the application of topic modelling.

LDA topic modelling is applied with standard 10 topics. Topic modelling is then analyzed by using the word cloud for each topic. Word cloud helps to visualize the words in topics. Further distances are calculated for each topic and hierarchal clustering is applied to explore the similarities among the topics.

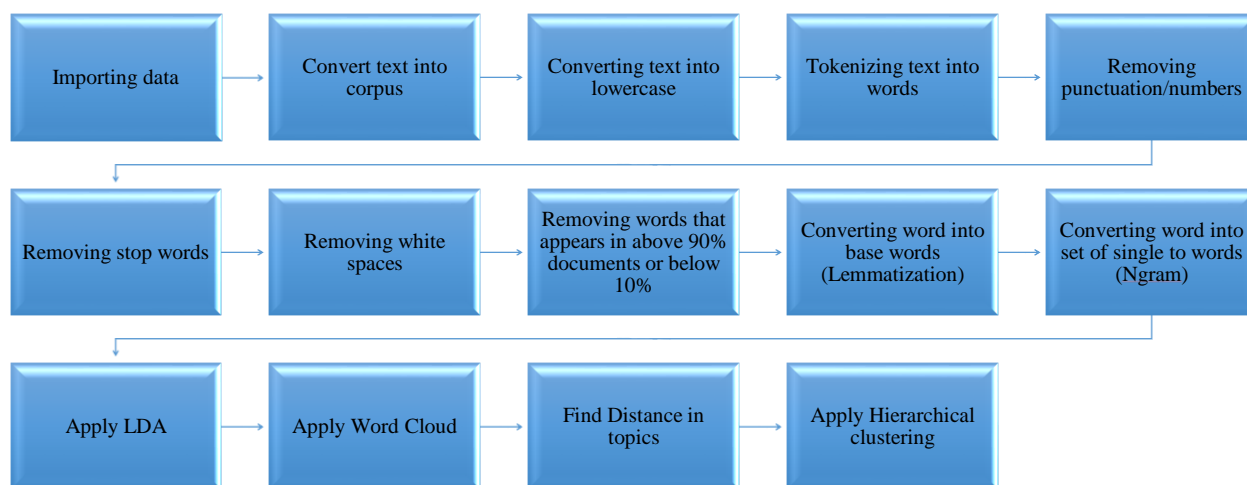


Figure 1: Preprocessing and topic modelling flow.

The topic labeling is subjective and based on topic background and theoretical knowledge of the research area. However, Guo et al (2017) and Lucini et al (2020) proposed that mechanism that

topic name should be based on top three terms that are also in line with other top 20 words. Topic name selection in this study is also based on top words that are associated with other majority words in the topic.

RESULTS

The complete model of topic modelling and analysis applied in this study is provided below

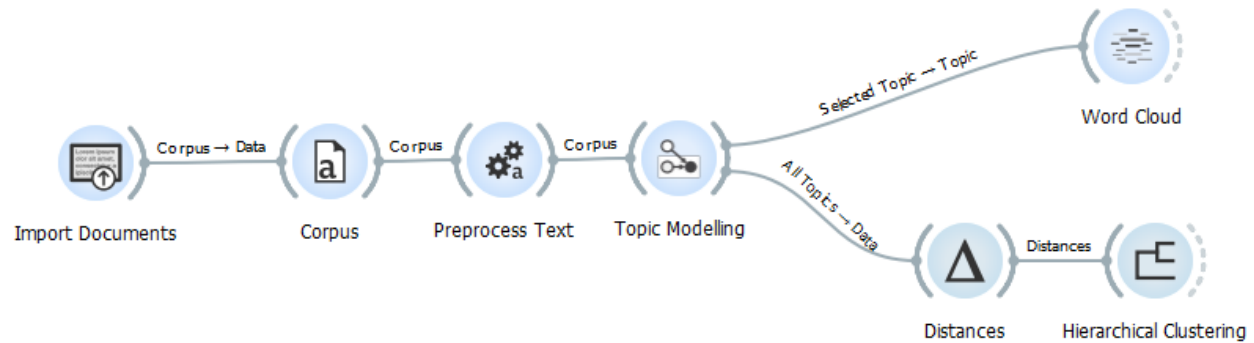


Figure 2: Model use for topic modelling in Orange

LDA Topic modelling is applied by using the above research model. The topic number is selected ten which is the default recommended for topic modelling in this software. The default topic number is selected to avoid biases. The following ten topics are identified

1. Environmental Responsibility
2. Digital Responsibility
3. Digital Skills
4. Digital Knowledge
5. Data Solutions
6. Quality Training
7. Quality Knowledge
8. Information Excellence
9. Software
10. Management Skills

Figure 3: List of identified topics

Discussion

Interestingly the first topic that emerges is environmental responsibility. All new development in the recent era is entirely focused to preserve the environment with industrial production. The adoption of industry 4.0 also fundamentally preserves the environment by improving efficiency (Kazancoglu et al, 2021). The quality professionals are also responsible to lead the organization in a way to preserve the environment

Digital responsibility is a new term that emerges in this study. Digital responsibility is a broader area and it includes how companies are collecting information, how to secure information as per privacy laws, and also how effectively use the information to increase productivity (Lobschat et al 2021). Therefore, quality professionals must also be digitally responsible.

Digital Skills are emerging as the third topic and quality professionals are now ranked on their digital skills (Hecker et al 2021) and how they can utilize the digital skills to optimize quality operations and provide more value in terms of real-time analysis and prescriptive analytics.

Digital Knowledge also appears as a separate topic in this study that indicates digital skills are not only enough for the quality professional but also digital knowledge on a broader domain. This will also provide the base for digital responsibility as well. Therefore, future quality professionals must also equip themselves with digital knowledge as well



Figure 4: Topic 1



Figure 5: Topic 2



Figure 6: Topic 3



Figure 7: Topic 4

The fifth topic is digital solutions. Digital Solution are now widely in practice to achieve the best results. Industry 4.0 is all about digital solutions (Ammar et al 2021) and quality 4.0 is also used digital solutions like computer vision and Artificial intelligence to control processes more effectively.



Figure 8: Topic 5

Quality Training is the sixth topic in this study. Training is an ongoing process in quality. Training of employees on new tools and techniques provides the best results in terms of customer satisfaction (Shen & Tang 2018).

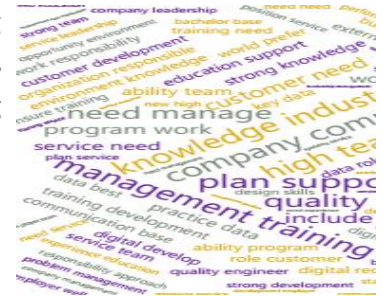


Figure 9: Topic 6

Quality knowledge also appears as a separate topic. Quality knowledge is upgraded day by day and quality knowledge will help professionals to understand the issues in more detail and devise the solution accordingly.



Figure 10: Topic 7

The next topic is information excellence. This is important because information excellence means how effectively information is utilized to meet customer requirements. Digital responsibility is also focused on effective utilization but is more oriented towards privacy. According to Radziwill (2018) quality, 4.0 is about information quality.



Figure 11: Topic 8

The 9th topic is software. It covers both dimensions. First, the software quality, and second that is more relevant to this context is the application of software in quality. Modern quality is all about the application of the latest software to collect, analyze and present the data. Additional application AI to analyze the take corrective action immediately in an automated fashion (Campbell et al, 2020).



Figure 12: Topic 9

Last is management skills. Management skills are still part of the game but updated as now data management skills are added and all management skills change from old methods to new methods like smart sheets.



Figure 13: Topic 10

The inter-topic similarity is also important to analyze how much topics are interrelated with each other. Hierarchical cluster is used to analyze the internal topic similarity. The dendrogram of the output is provided below.

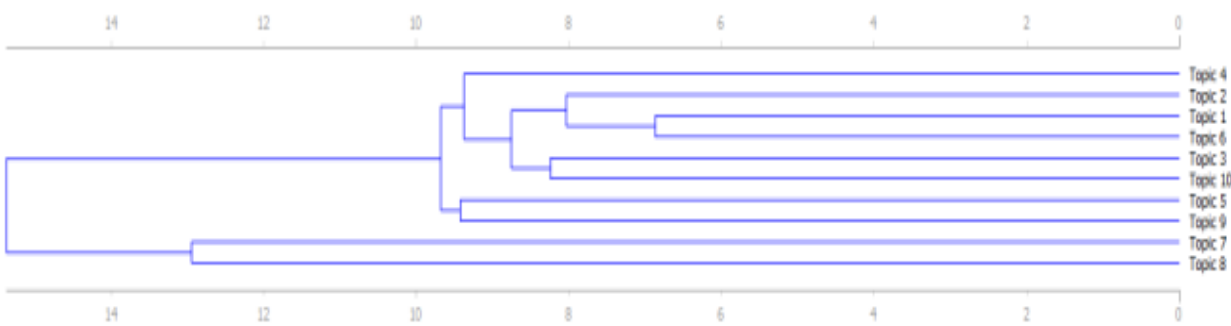


Figure 14: Denrogram for intertopic smiliarity

Environmental responsibility and quality training are most similarly based on their height in the dendrogram. After this, these two topics are collectively similar to digital responsibility. These similarities are logical are quality training must include the environmental and digital responsibilities so the quality professional can perform accordingly.

Management skills and Digital skills are the next most similar topics in this research. This indicates that management skills and digital skills are closely interrelated and therefore digital skills must be included in management skills training because now digital skills are pre-requisite to manage any business.

Data Solutions and software has also similarities and definitely, it is quite logical because data solutions are software that enables the management of the data effectively and enable the utilize data for better productivity and customer satisfaction.

Quality knowledge and information excellence have the least similarity in this research. Quality knowledge is a broader domain and covers tools and techniques to optimize the process while information is specific to accessing and utilization of the best information.

Additionally, it is observed that quality knowledge and information excellence are collectively least similar to all other topics as well.

CONCLUSION

Job advertisements listed on the worlds largest postal are selected against keywords quality 4.0. digital quality, smart quality, and quality in industry 4.0. LDA Topic Modeling is applied by using Orange 3.28.0. The results identify the key ten topics in the job descriptions are Environmental Responsibility, Digital Responsibility, Digital Skills, Digital Knowledge, Data Solution, Quality Training, Quality Knowledge, information excellence, software, and management skills.

This research provides very important insights. The first topic is environmental responsibility that indicates that modern quality jobs are more oriented towards environmental preservation. Industry 4.0 is also fundamental to preserving the environment by optimizing the inter and intra organizational processes.

Digital Responsibility is also one of the key contributions of this research. Social responsibility is considered at the organization level but not previously specifically linked with quality. However, Digital responsibility was not highlighted previously. This research identifies it as an important topic. Quality is now transformed from traditional quality to quality 4.0 and digital responsibility is one of the major addition to quality.

Digital responsibility is about the collection of the data, the means of collection and compliance of the privacy and legal guidelines, the development of data storage to ensure the safety of the data, and finding the best option to utilize the data to meet the customer requirement effectively.

Digital Skills, Digital Knowledge, software, and Data Solution are also identified as key topics in this research. Digital Skills, software, and Data solution are part of quality 4.0 but interestingly digital knowledge is also identified as that indicate that quality professional not only focus on specific digital skills but also acquire digital knowledge to effectively understand and implement modern quality management. Quality Knowledge, Quality Training, and Management skills are also related to the context

Information Excellence is also identified. This is in line with the information quality concept. The modern quality is all about what to know, why to know and how to know. Information excellence is another key contribution of this study along with digital responsibility.

Environmental responsibility, digital responsibility, and quality training have higher similarities in comparison to the other topics. Quality knowledge and information excellence have the least similarity in this research and collectively have the least similarity with all other topics as well.

The job advertisement provides insights about the organization intension and plans and analysis of the job description identified that now the organizations are shifting their intention from traditional quality to quality 4.0 because the no topic identified in this research related to traditional quality like process improvement, process control, etc. therefore, this study concluded that traditional quality is transformed into quality 4.0.

This research is based on job advertisements to determine what organizations are looking and topic modelling is applied to identify the key area. This research points out digital responsibility and information excellence. Future research further explores these contexts in case studies and a quantitative approach. Additionally, also conduct this study in different sectors and countries to identify the global applicability.

REFERENCE

1. Ammar, M., Haleem, A., Javaid, M., Walia, R., & Bahl, S. (2021). Improving material quality management and manufacturing organizations system through Industry 4.0 technologies. *Materials Today: Proceedings*, 45, 5089-5096.
2. Arora, N., & Gupta, S. (2020). Total Quality Management for Employee Engagement: A Study. *Test Engineering and Management*, 82, 12769-86.
3. Asmussen, C. B., & Møller, C. (2019). Smart literature review: a practical topic modelling approach to exploratory literature review. *Journal of Big Data*, 6(1), 1-18.
4. ASQ (2018). Industry and quality 4.0: bringing them together. *Quality Magazine*, Retrieved July 08, 2021, from: <https://www.qualitymag.com/articles/95011-industry-and-quality-40-bringing-them-together>
5. Campbell, C., Sands, S., Ferraro, C., Tsao, H. Y. J., & Mavrommatis, A. (2020). From data to action: How marketers can leverage AI. *Business Horizons*, 63(2), 227-243.
6. Durana, P., Kral, P., Stehel, V., Lazaroiu, G., & Sroka, W. (2019). Quality Culture of Manufacturing Enterprises: A possible way to adaptation to Industry 4.0. *Social Sciences*, 8(4), 124.
7. Feldman, S. (2005). Quality Assurance: Much More than Testing: Good QA is not only about technology, but also methods and approaches. *Queue*, 3(1), 26-29.
8. Gara, J. A., Zakaria, R. B., Aminudin, E., Adzar, J. A., & Yousif, O. S. (2021). The Development of Real-Time Integrated Dashboard: An Overview for Road Construction Work Progress Monitoring. *Journal of Hunan University Natural Sciences*, 48(5).

9. Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: Tourist satisfaction analysis using latent dirichlet allocation. *Tourism management*, 59, 467-483.
10. Haynes, R. B., & Walker, C. J. (1987). Computer-aided quality assurance: a critical appraisal. *Archives of internal medicine*, 147(7), 1297-1301.
11. Hecker, I., Spaulding, S., & Kuehn, D. (2021). Digital Skills and Older Workers. Urban Institute, September, 14.
12. Jacob, D. (2017). Top 4 Reasons to Update to Quality 4.0. [online] blog.Insresearch.com. Retrieved July 08, 2021, from: <https://blog.Insresearch.com/top-4-reasons-to-update-to-quality-4.0>
13. Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia Tools and Applications*, 78(11), 15169-15211.
14. Kazancoglu, Y., Sezer, M. D., Ozkan-Ozen, Y. D., Mangla, S. K., & Kumar, A. (2021). Industry 4.0 impacts on responsible environmental and societal management in the family business. *Technological Forecasting and Social Change*, 173, 121108.
15. Kolini, F., & Janczewski, L. (2017). Clustering and topic modelling: A new approach for analysis of national cyber security strategies.
16. Kim, N., Lee, D., Choi, H., & Wong, W. X. S. (2017). Investigations on techniques and applications of text analytics. *The Journal of Korean Institute of Communications and Information Sciences*, 42(2), 471-492.
17. Kodama, F. (2018). Learning mode and strategic concept for the 4th Industrial Revolution. *Journal of Open Innovation: Technology, Market, and Complexity*, 4(3), 32.
18. Lobschat, L., Mueller, B., Eggers, F., Brandimarte, L., Diefenbach, S., Kroschke, M., & Wirtz, J. (2021). Corporate digital responsibility. *Journal of Business Research*, 122, 875-888.
19. Lucini, F. R., Tonetto, L. M., Fogliatto, F. S., & Anzanello, M. J. (2020). Text mining approach to explore dimensions of airline customer satisfaction using online customer reviews. *Journal of Air Transport Management*, 83, 101760.
20. Negara, E. S., Triadi, D., & Andryani, R. (2019). Topic Modelling Twitter Data with Latent Dirichlet Allocation Method. In 2019 International Conference on Electrical Engineering and Computer Science (ICECOS) (pp. 386-390). IEEE.
21. Nica, E., & Stehel, V. (2021). Internet of Things sensing networks, artificial intelligence-based decision-making algorithms, and real-time process monitoring in sustainable Industry 4.0. *Journal of Self-Governance and Management Economics*, 9(3), 35-47.
22. Polner, E. (2021, December 21). The 10 best job search websites of 2022. The Balance Careers. Retrieved January 18, 2022, from <https://www.thebalancecareers.com/top-best-job-websites-2064080#:~:text=Indeed%20was%20founded%20in%202004,job%20listings%20added%20every%20second.>

23. Radziwill, N. (2018). Let's get digital. *Quality Progress*, 51(10), 24-29.
24. Rubinger, L., Gazendam, A., Ekhtiari, S., Nucci, N., Payne, A., Johal, H., ... & Bhandari, M. (2020). Maximizing virtual meetings and conferences: a review of best practices. *International orthopaedics*, 44, 1461-1466.
25. Shen, J., & Tang, C. (2018). How does training improve customer service quality? The roles of transfer of training and job satisfaction. *European management journal*, 36(6), 708-716.
26. Sony, M., Antony, J., Douglas, J. A., & McDermott, O. (2021). Motivations, barriers and readiness factors for Quality 4.0 implementation: an exploratory study. *The TQM Journal*.
27. Srinivasan, A., & Kurey, B. (2014). Creating a culture of quality. *Harvard Business Review*, 92(4), 23-25.
28. Tong, Z., & Zhang, H. (2016). A text mining research based on LDA topic modelling. In *International Conference on Computer Science, Engineering and Information Technology* (pp. 201-210).
29. Tsutsui, W. M. (2019). History of Japanese Labor and Production Management. In *Oxford Research Encyclopedia of Business and Management*.
30. Villalba-Diez, J., Schmidt, D., Gevers, R., Ordieres-Meré, J., Buchwitz, M., & Wellbrock, W. (2019). Deep learning for industrial computer vision quality control in the printing industry 4.0. *Sensors*, 19(18), 3987.
31. Wang, M., Wang, C. C., Sepasgozar, S., & Zlatanova, S. (2020). A systematic review of digital technology adoption in off-site construction: Current status and future direction towards industry 4.0. *Buildings*, 10(11), 204.
32. Watson, Gregory H; Sörqvist, Lars; Keim, Elizabeth M; (2018). *The Journal for Quality and Participation*; 40 (4), 4-10
33. Zonnenshain, A., & Kenett, R. S. (2020). Quality 4.0—the challenging future of quality engineering. *Quality Engineering*, 32(4), 614-626.