SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2020
ORGANIZATION AND MANAGEMENT SERIES NO. 142

IDENTIFICATION OF TECHNOLOGIES IN INDUSTRY 4.0 WITH THE USE OF TEXT MINING

Dariusz ZDONEK

Silesian University of Technology, Faculty of Organization and Management, Zabrze; Dariusz.Zdonek@polsl.pl, ORCID: 0000-0002-6190-9643

Purpose: The objective of this paper is to identify leading technologies in Industry 4.0.
Design/methodology/approach: The identification was made with the use of text mining to explore the scientific texts in this field. Assumptions of own iterative method for analyzing scientific texts were proposed, with the use of R language, tokenization, lemmatization, n-grams and correspondence analysis. The assumptions of the proposed method were used to analyze the 40 most often quoted articles indexed in the Web of Science.
Findings: On the basis of the obtained results, 4 leading technologies were identified. These are Cloud Computing, Internet of Things, Cyber-physical System and Big Data.
Originality/value: The article proposes an original method of identifying the leading technologies used in Industry 4.0. The proposed method is based on text mining and correspondence analysis.

Keywords: Text mining, industry 4.0, information and communication technology, scientific papers.

Category of the paper: Research paper.

1. Introduction

The idea of "Industry 4.0" is linked with a vision of the world, that in the future will reach development level defined as the fourth industrial revolution. The ever present digitalization enforced by the globalization of markets results in increasing changes in the global economy. This influences not only the users of services but also the structure of the market of service providers and manufacturers. The changes taking place on the market of services and production can be seen even today, and it is expected that they will increase in the future. According to the Digital Vortex report (Bradley et al., 2015), the most radical changes can take place in data-driven industries, such as new technologies, media and entertainment, financial services, telecommunication, as well as retail. At the same time, great development opportunities for companies, that will be able to adjust to the changing environmental conditions and will be able to use digital technologies to develop their products and services, are being emphasized. Therefore, more and more managers pay attention to the newly introduced IT services and solutions that directly or indirectly influence the market. A source of information can be fairs, industry papers, popular science articles, business meetings and conferences (such as: Apple Special Events, I/O '19 etc.), where specialists, manufacturers, service providers, scientists and business practitioners share information, experience and create visions for the development of the whole field. Around the world, there are more and more conferences linked directly or indirectly with Industry 4.0 (for example Future Manufacturing Technologies, Industries 4.0, Smart Industry Expo etc.). During these conferences and own research, there are more and more articles being published in scientific papers.

The aim of this article is to identify leading technologies linked with Industry 4.0. The research covered the analysis of definitions, schemes available on the web and 40 selected articles indexed in the Web of Science, that were displayed for the search "Industry 4.0". For the analysis of the text, an own method in the area of text mining was used.

2. Definitions and schemes of Industry 4.0.

There are many Industry 4.0 definitions. For instance, Trappey et al. (2016) defined Industry 4.0 as a general concept enabling manufacturing with the elements of tactical intelligence using techniques and technologies, such as Internet of things, cloud computing and big data. For Schumacher, Erol, and Sihn (2016) Industry 4.0 refers to recent technological advances, where the Internet and supporting technologies (e.g. embedded systems) serve as a backbone to integrating physical objects, human actors, intelligent machines, product lines and processes across organizational boundaries, to form a new kind of intelligent, networked and agile value chain. Whereas Kohler and Weisz defined Industry 4.0 as a new approach for controlling production processes by providing real time synchronization of flows and by enabling the unitary and customized fabrication of products. (Kohler et al. 2016; Moeuf, 2018).

In all of these definitions, the concept of Industry 4.0 is based on the use of new technologies, that should facilitate the flow of information, control and adjustment of operations in the whole production system in real time. The interest in Industry 4.0 is still growing. This is proven by the trend graphs in Google Trends (Figure 1), as well as by the number of publications indexed in the Web of Science (Figure 2).

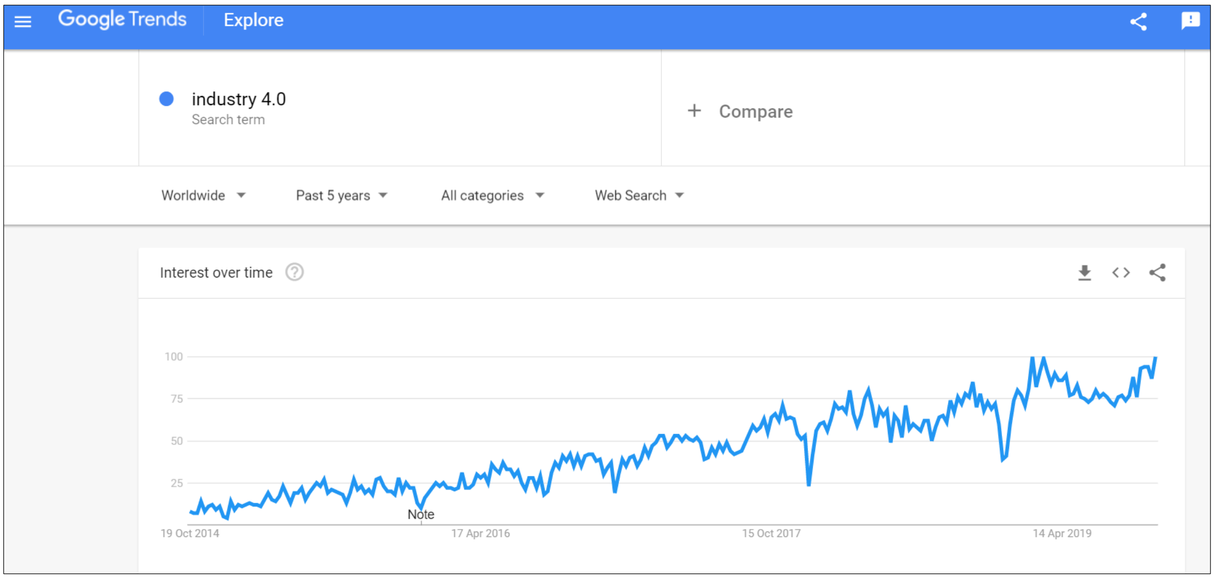

Figure 1. Google Trends 2014-2019 for question “Industry 4.0”.

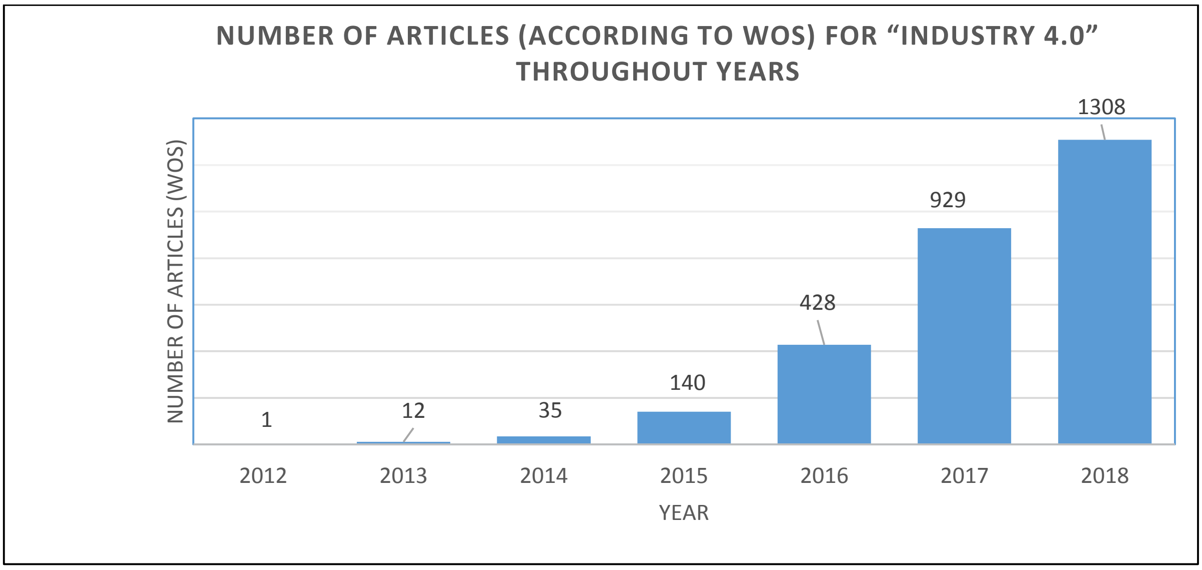

Figure 2. Number of articles (according to WoS) for “Industry 4.0” 2012-2018, 29.10.2019. Own elaboration based on https://apps.webofknowledge.com.

This translates into even greater number of definitions, terms and schemes defining the Industry 4.0. Since in the area of Industry 4.0 technologies play a crucial role, it was decided to distinguish them on the basis of schemes available on the Internet. For the search “Industry 4.0 technologies” a large number of various schemes presenting technologies in the area of Industry 4.0 was displayed. (Figure 3).

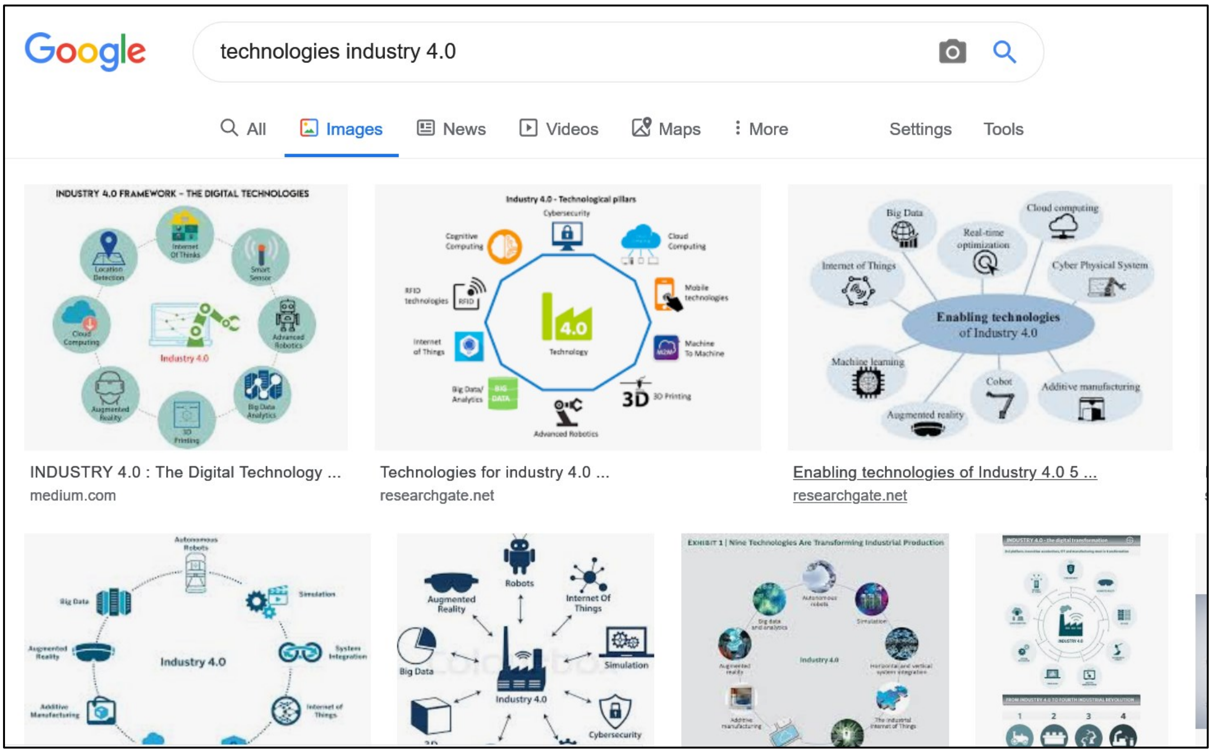

Figure 3. Images for "technologies industry 4.0".

It was decided to analyze 6 randomly selected schemes and then identify these technologies that were mentioned there. The results sorted by frequency are presented in Table 1.

Table 1.
Analysis of schemes concerning technologies in Industry 4.0

ID.	Technology	Fig. 1	Fig. 2	Fig. 3	Fig. 4	Fig. 5	Fig. 6
1	Cloud Computing	x	x	x	x		x
2	Augmented Reality	x			x	x	x
3	Internet of Things	x	x	x	x		
4	Cybersecurity		x		x	x	
5	Additive manufacturing			x	x	x	
6	Big Data Analytics	x	x			x	
7	3D Printing	x	x				x
8	Big Data			x	x		x
9	Advanced Robotics	x	x				
10	Cyber Physical System			x			x
11	Autonomous Robots				x	x	
12	Simulation				x	x	
13	System integration				x		
14	Horizontal and vertical system integration					x	
15	The industrial Internet of Things					x	
16	The cloud					x	
17	System Security						X
18	Humanoid Robots						X
19	Location Detection	x					
20	Machine learning			x			
21	Cobot			x			
22	Smart Sensor	x					

Cont. table 1.

23	Mobile technologies		x				
24	Machine to Machine		x				
25	RFID technologies		x				
26	Cognitive Computing		x				
27	Real-time optimization			x			

Fig. 1 – https://medium.com/@winix/industry-4-0-the-digital-technology-transformation-b23ba02a7dd2.
Fig. 2 – https://www.researchgate.net/figure/Technologies-for-industry-40_fig1_319944621.
Fig. 3 – https://www.researchgate.net/figure/Enabling-technologies-of-Industry-40-5-ASSEMBLY-SYSTEM-40_fig1_320496225.
Fig. 4 – https://aethon.com/mobile-robots-and-industry4-0/.
Fig. 5 – https://www.robotlab.com/blog/the-nine-technologies-of-industry-4.0.
Fig. 6 – https://iot-analytics.com/industrial-technology-trends-industry-40-patents-12x/.

Source: own elaboration.

It turned out, that – similarly as in the case of definition of Industry 4.0 – there is currently no single scheme of technologies that is linked with Industry 4.0. The authors of the publications and schemes often use different names for the same technologies. The lack of unified templates is, on the one hand, troublesome for the reader, who wants to understand a given term and, on the other hand, this is a natural process of science and research development in a new field.

Exemplary analyzed schemes presenting the technologies linked with Industry 4.0 are presented in Figure 4.

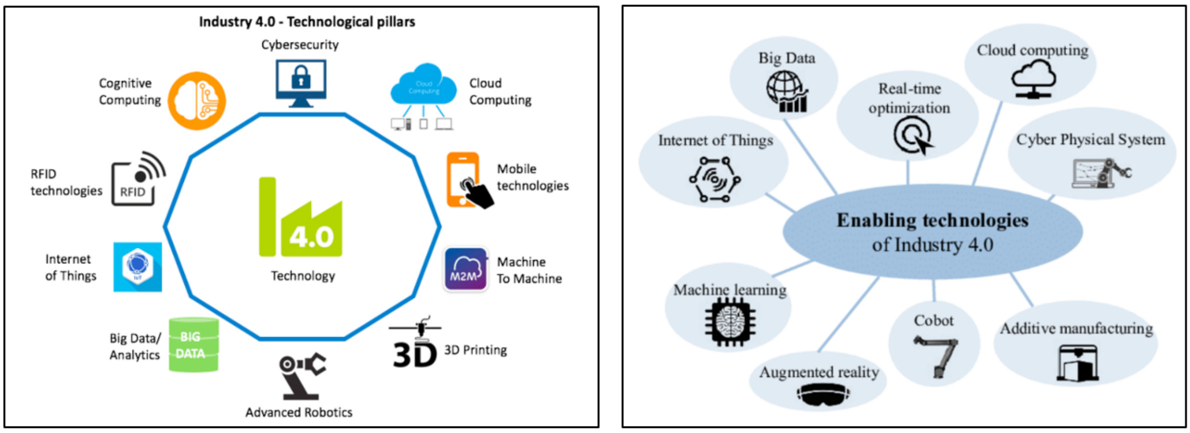

Figure 4. Technological pillars and enabling technologies of Industry 4.0. Own elaboration based on (Saturno et al. 2017, Bortolini et al. 2017).

The analysis of definitions and schemes is one of the manners of identification of technologies in the area of Industry 4.0. Despite the fact that it is a time-consuming process, with a large number of analyzed schemes a list of the most frequently named technologies can be obtained. A flaw of this method is the fact that schemes are often copied by many authors and they do not take into account the meaning and development of a given technology in the field of Industry 4.0. It was assumed, that the more changes take place in a given technology and its use is the area of research within Industry 4.0, the more articles there are. Therefore, it was decided that a complementation of this analysis can be a study of scientific texts and verification how frequently certain words and phrases associated with technologies appear in them.

3. Text mining in the analysis of scientific texts on Industry 4.0

In order to carry out the analysis of text documents, a text mining method was used, that is about automatic identification of useful information from a large number of unstructured text data. It uses the process of tokenization, elimination of stop words, stemming and pop-tag (Fan et. al., 2006; Vijayarani et al. 2015; Silge, & Robinson, 2017). The initial analysis was carried out on a set of 40 academic texts concerning the Industry 4.0. 19 of the most frequently quoted articles from 2014 (from Web of Science Core Collection) and 21 of the most frequently quoted from 2018 were downloaded (Table 1). The documents were downloaded in PDF format. The research was carried out in October 2019.

Table 2.
The most frequently quoted scientific papers concerning "Industry 4.0" acc. to Web of Science published in 2014 and in 2018

Id	Article	Times Cited
1	Lasi, H., Fettke, P., Kemper, H.G., Feld, T., & Hoffmann, M. (2014). Industry 4.0.	402
2	Lee, J., Kao, H.A., & Yang, S. (2014). Service innovation and smart analytics for industry 4.0 and big data environment.	361
3	Monostori, L. (2014). Cyber-physical production systems: Roots, expectations and R&D challenges.	297
4	Shrouf, F., Ordieres, J., & Miragliotta, G. (2014, December). Smart factories in Industry 4.0: A review of the concept and of energy management approached in production based on the Internet of Things paradigm.	151
5	Gorecky, D., Schmitt, M., Loskyll, M., & Zühlke, D. (2014, July). Human-machine-interaction in the industry 4.0 era.	99
6	Jazdi, N. (2014, May). Cyber physical systems in the context of Industry 4.0	90
7	Varghese, A., & Tandur, D. (2014, November). Wireless requirements and challenges in Industry 4.0.	51
8	Ungurean, I., Gaitan, N.C., & Gaitan, V.G. (2014, May). An IoT architecture for things from industrial environment.	35
9	Vogel-Heuser, B., Diedrich, C., Pantförder, D., & Göhner, P. (2014, July). Coupling heterogeneous production systems by a multi-agent based cyber-physical production system.	32
10	Dombrowski, U., & Wagner, T. (2014). Mental strain as field of action in the 4th industrial revolution.	29
11	Schuh, G., Potente, T., Varandani, R., & Schmitz, T. (2014). Global Footprint Design based on genetic algorithms – An "Industry 4.0" perspective.	29
12	Schuh, G., Potente, T., Varandani, R., Hausberg, C., & Fränken, B. (2014). Collaboration moves productivity to the next level.	18
13	Stark, R., Grosser, H., Beckmann-Dobrev, B., Kind, S., & INPIKO Collaboration. (2014). Advanced technologies in life cycle engineering.	11
14	Sauer, O. (2014). Information Technology for the Factory of the Future – State of the Art and Need for Action.	7
15	Constantinescu, C.L., Francalanza, E., Matarazzo, D., & Balkan, O. (2014). Information support and interactive planning in the Digital Factory: Approach and industry-driven evaluation.	4
16	Thoben, K.D., Busse, M., Denkena, B., & Gausemeier, J. (2014). System-integrated Intelligence – New Challenges for Product and Production Engineering in the Context of Industry 4.0.	4
17	Gath, M., Herzog, O., & Edelkamp, S. (2014, October). Autonomous and flexible multiagent systems enhance transport logistics.	3

Cont. table 2.

18	Jaschke, S. (2014, December). Mobile learning applications for technical vocational and engineering education: The use of competence snippets in laboratory courses and industry 4.0	3
19	Bauer, W., Ganschar, O., Pokorni, B., & Schlund, S. (2014). Concept of failure management assistance system for the reaction on unforeseeable events during the ramp-up.	2
20	Xu, L.D., Xu, E.L., & Li, L. (2018). Industry 4.0: state of the art and future trends.	95
21	Li, L. (2018). China's manufacturing locus in 2025: With a comparison of "Made-in-China 2025" and "Industry 4.0".	61
22	Chen, B., Wan, J., Shu, L., Li, P., Mukherjee, M., & Yin, B. (2018). Smart factory of industry 4.0: Key technologies, application case, and challenges.	55
23	Qi, Q., & Tao, F. (2018). Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison.	46
24	Moeuf, A., Pellerin, R., Lamouri, S., Tamayo-Giraldo, S., & Barbaray, R. (2018). The industrial management of SMEs in the era of Industry 4.0.	43
25	Müller, J.M., Kiel, D., & Voigt, K.I. (2018). What drives the implementation of Industry 4.0? The role of opportunities and challenges in the context of sustainability.	42
26	Müller, J.M., Buliga, O., & Voigt, K.I. (2018). Fortune favors the prepared: How SMEs approach business model innovations in Industry 4.0.	41
27	Tseng, M.L., Tan, R.R., Chiu, A.S., Chien, C.F., & Kuo, T.C. (2018). Circular economy meets industry 4.0: can big data drive industrial symbiosis?	38
28	Yin, Y., Stecke, K.E., & Li, D. (2018). The evolution of production systems from Industry 2.0 through Industry 4.0.	35
29	Fraga-Lamas, P., Fernández-Caramés, T.M., Blanco-Novoa, Ó., & Vilar-Montesinos, M.A. (2018). A review on industrial augmented reality systems for the industry 4.0 shipyard.	33
30	de Sousa Jabbour, A.B.L., Jabbour, C.J.C., Foropon, C., & Godinho Filho, M. (2018). When titans meet – Can industry 4.0 revolutionize the environmentally-sustainable manufacturing wave? The role of critical success factors.	32
31	Zheng, P., Sang, Z., Zhong, R.Y., Liu, Y., Liu, C., Mubarok, K., ... & Xu, X. (2018). Smart manufacturing systems for Industry 4.0: Conceptual framework, scenarios, and future perspectives.	32
32	Khaqqi, K.N., Sikorski, J.J., Hadinoto, K., & Kraft, M. (2018). Incorporating seller/buyer reputation-based system in blockchain-enabled emission trading application.	30
33	de Sousa Jabbour, A.B.L., Jabbour, C.J.C., Godinho Filho, M., & Roubaud, D. (2018). Industry 4.0 and the circular economy: a proposed research agenda and original roadmap for sustainable operations.	28
34	Jiang, Y., Yin, S., & Kaynak, O. (2018). Data-driven monitoring and safety control of industrial cyber-physical systems: Basics and beyond.	27
35	Tortorella, G.L., & Fettermann, D. (2018). Implementation of Industry 4.0 and lean production in Brazilian manufacturing companies. International Journal of Production Research,	27
36	Sisinni, E., Saifullah, A., Han, S., Jennehag, U., & Gidlund, M. (2018). Industrial internet of things: Challenges, opportunities, and directions.	26
37	Fu, Y., Ding, J., Wang, H., & Wang, J. (2018). Two-objective stochastic flow-shop scheduling with deteriorating and learning effect in Industry 4.0-based manufacturing system.	26
38	Li, X., Li, D., Wan, J., Liu, C., & Imran, M. (2018). Adaptive transmission optimization in SDN-based industrial Internet of Things with edge computing.	26
39	Sung, T.K. (2018). Industry 4.0: a Korea perspective.	24
40	Aazam, M., Zeadally, S., & Harras, K.A. (2018). Deploying fog computing in industrial internet of things and industry 4.0.	23

Note. Own elaboration based on Web of Science (https://apps.webofknowledge.com, 2019.11.15).

3.1. Methods

Within the conducted research, 5 characteristic stages of proceeding were distinguished (Figure 5).

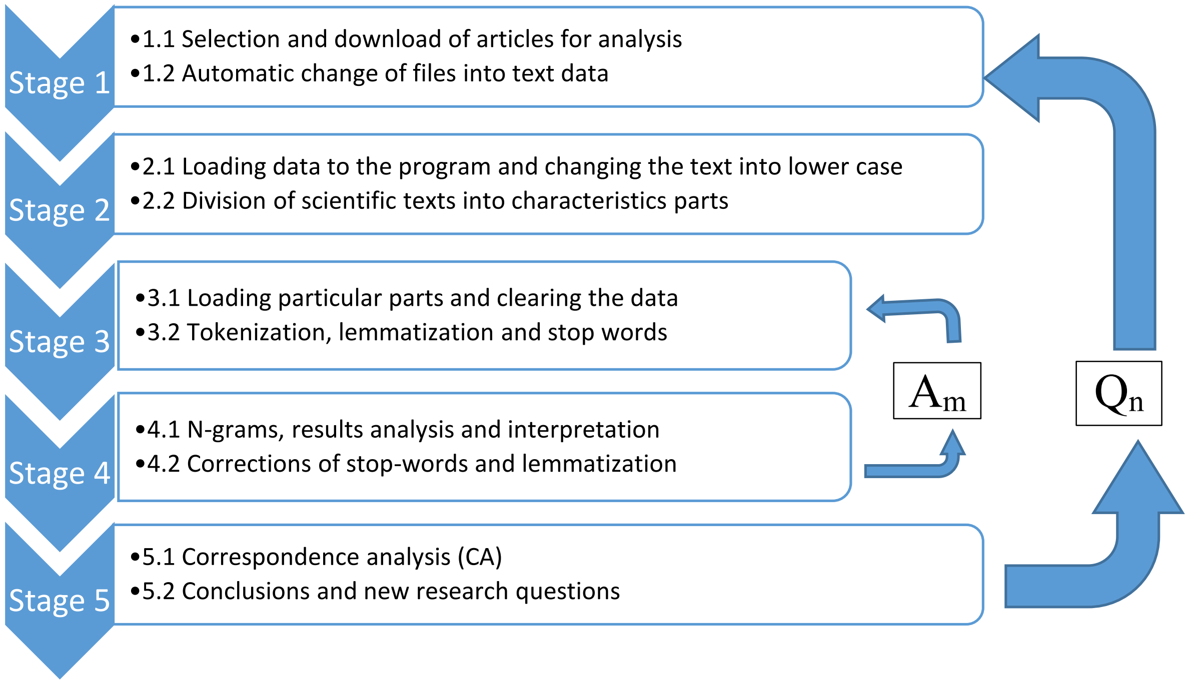

Figure 5. Research stages – assumptions for Simple Text Mining n-iteration (STMn). Own elaboration.

The first stage was the selection and download of articles for analysis and their conversion into text files. For that purpose, https://apps.webofknowledge.com website made it possible to search for indexed scientific papers in WoS database and https://pdftotext.com/pl/ tool available online was used for automatic change of file format into txt. Thanks to the extended possibilities of searching articles on the topic of Industry 4.0, 19 of the most frequently quoted articles from 2014 and 21 from 2018 were downloaded in the PDF format. Next, these files were converted into txt.

The second stage was about loading all collected data in text files into the analytical tool. For that purpose, RStudio software was used.

Next, all characters were converted into lowercase letters and it was checked if the articles contain characteristic single word "references". For that purpose, relevant scripts in R language were used. In case of files, where more or less than one "references" word was present, a manual correction on text files had to be made and the data had to be reloaded to RStudio. As a result, it was possible to divide data into two characteristic areas (that is content and references).

The third stage was about loading to separate variables contents and the references of all articles and about clearing the text of unnecessary words, phrases, characters that were in the text or were incorrectly saved during the conversion of PDF files into TXT.

After initial clearing, the text data was subject to tokenization, lemmatization and stop words. Tokenization is about dividing data into single words. Lemmatization is about changing the words in plural into words in singular (for example: industries into industry, systems into system, technologies into technology etc.) and about reversing the words to their main root (for example: computing into compute, given into give etc.). Stop words is about deleting words that do not have any meaning in analytics (for example: the, a, an, in, is, at, and etc.).

At this stage, POS-tag was ignored on purpose (Nakagawa, 2007). It forms a part of NLP and concerns marking the words with an appropriate part of the speech (for example: noun, adjective, verb etc.). It was decided that it is not necessary for the purpose of this research.

The fourth stage was about creating frequency statistics of the used single, double or triple words in the text with the use of n-grams. Moreover, an analysis of the results and correction of stop words list was made. The corrections were about adding additional characters or words to stop words. Therefore, in the proposed methodology, there is an iterative feedback loop A_n.

The last stage was about analyzing the obtained information and creating the new knowledge on the analyzed documents with the use of selected method of data analysis. In the described method, the use of words statistics was proposed with the use of n-grams and correspondence analysis.

Once the results are obtained, quite often additional research questions emerge, that require new analysis and correction of assumptions. Therefore, Q_n loop returning to Stage 1 was introduced. It was assumed, that the analysis of scientific text documents is a constant and iterative process. With every stage, it is possible to obtain additional or more detailed information.

3.2. Results and conclusions

Using the assumptions and stages presented in the above chapter, a research covering 40 documents on Industry 4.0 was carried out. In the first Q_1 iteration, the list of the most frequently used words and phrases on Industry 4.0 in the analyzed documents was obtained. In total, 8,782 words were obtained, 3,846 (approx. 44%) of which were used only once in all documents. The most frequently used words were industry, system, data, manufacture, production, process, technology etc. (Figure 6).

Word	Freq.
industry	1917
system	1601
data	1413
manufacture	1245
production	1033
process	1009
technology	929
product	814
industrial	734
information	697
time	653
smart	634
base	618
machine	556
model	534
company	515
network	491

Figure 6. Single words. Own elaboration with the use of RStudio and https://www.wordclouds.com.

Taking into account 2-grams (two-word words), the most frequently used were real time, smart factory, cyber physical etc. (Figure 7).

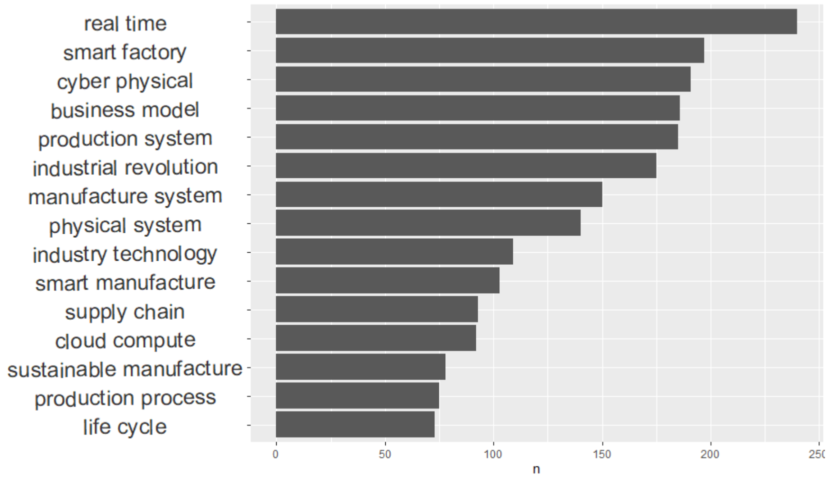

Figure 7. Double words (2-grams). Own elaboration.

In case of 3-grams, 101,496 combinations of words were obtained, of which 964,330 (approx. 95%) were present only once. After deleting the words occurring often, but in single articles, and deleting the names of periodicals, an interesting list of phrases linked with Industry 4.0 was obtained. It can be noticed that they describe the Industry 4.0 very well. Table 3.

Table 3.
Phrases (3-grams)

Words	Freq.
cyber physical system	133
real time data	43
4th industrial revolution	39
forecast social change	39
business model innovation	34
international production research	32
smart manufacture system	31
critical success factor	23
physical production system	23
product life cycle	21
production plan control	19
competitiveness future viability	18
information communication technology	18
organizational production fit	18
real time information	18
employee qualification acceptance	16
monitor fault diagnosis	16
multi agent system	16

The last stage of the analysis with the use of the proposed method was to apply simple statistics of words and correspondence analysis to group documents and technologies that they describe. It was a simplified research and the results for such a small group of documents differed depending on the words that were selected as variables. One of such analyses is presented in Table 4, 5 and Figure 8, where the number of keywords linked with technologies in selected articles was calculated.

Table 4.
Correspondence analysis Article 2014

	Article 2014																			
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**Sum**
augment	1	0	0	0	4	0	0	0	0	0	0	0	2	0	0	0	0	0	0	7
cloud	0	3	2	4	0	13	2	0	10	1	0	0	5	0	1	0	1	0	1	43
print	1	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	2	0	8
iot	0	1	1	49	0	0	2	12	0	0	0	0	0	0	0	0	0	0	0	65
physical	8	11	15	3	13	6	5	5	17	7	0	32	7	1	1	4	1	8	2	146
smart	4	11	3	70	1	7	2	9	1	2	0	2	3	1	3	0	1	1	0	121
security	0	0	1	0	2	2	2	3	1	0	0	0	0	3	0	0	0	1	1	16
big	0	21	4	6	0	1	0	0	2	2	1	0	5	0	1	0	0	3	0	46
simulation	1	1	1	2	1	0	0	0	2	0	1	7	1	0	14	0	2	6	0	39
Sum	16	50	30	138	26	35	20	37	42	22	13	53	41	19	35	20	22	39	23	491

Table 5.
Correspondence analysis Article 2018

	Article 2018																				
	20	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**Sum**
Augment	0	0	0	3	1	0	1	0	0	19	0	2	0	0	0	0	0	0	0	3	0	29
Cloud	25	1	21	5	39	1	3	1	10	18	7	36	0	18	1	3	3	1	1	2	8	204
Print	0	2	0	1	1	0	0	0	29	2	1	4	0	0	0	3	0	0	0	2	0	45
Iot	84	5	7	7	13	0	6	2	10	3	0	12	0	2	4	1	47	0	16	2	16	237
Physical	36	9	18	43	10	3	10	1	16	10	7	24	0	11	14	2	11	2	0	10	2	239
Smart	31	3	92	40	10	10	17	0	34	13	2	138	8	9	28	2	18	3	6	4	39	507
Security	21	1	9	1	7	4	5	0	0	3	0	1	1	0	3	0	19	0	0	7	9	91
Big	2	3	23	120	11	0	2	5	8	1	2	15	0	5	2	4	2	1	3	3	4	216
simulation	0	0	0	13	21	2	1	0	0	1	1	1	0	1	0	1	0	5	3	0	0	50
Sum	219	45	192	256	137	45	71	36	135	99	50	264	41	79	86	51	136	49	67	72	118	1618

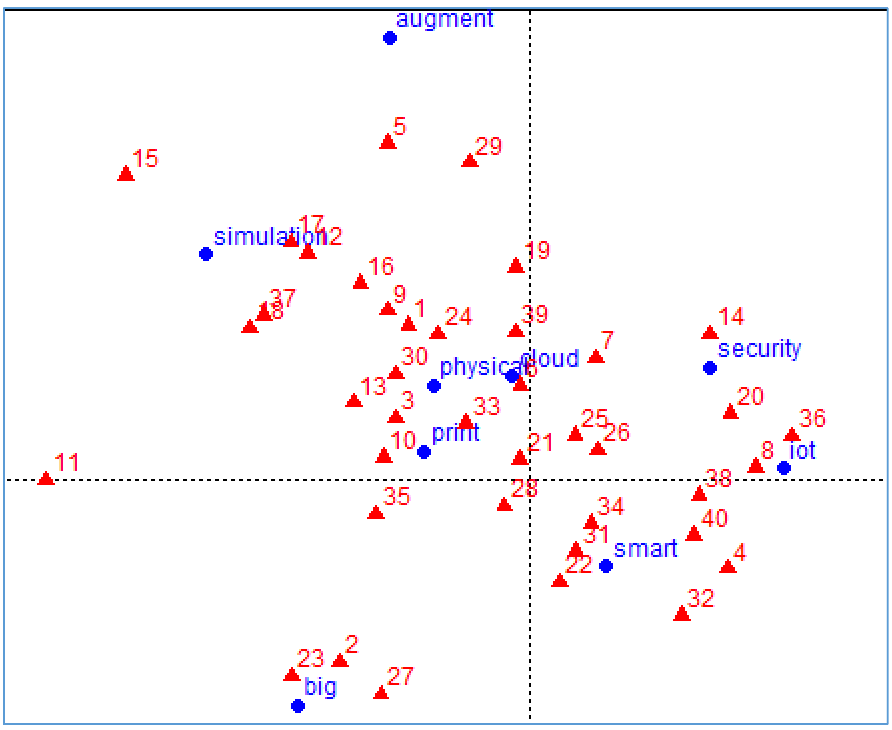

Figure 8. Correspondence analysis.

From the initial analysis, it can be noticed that from the texts subject to analysis, the most comprehensive in terms of IoT (Internet of things) are articles no. 36 and 8. Articles 2, 23 and 27 discuss the issues concerning big data. Articles 22, 31 and 34 are about smart and article

6 is about cloud computing. Article 11 is the least concerned with the studied technologies and articles no. 14 and 20 tackle security.

The results obtained with the analysis of schemes and text mining in the scope of technologies contain many similar terms. For example, cloud computing, big data, IoT, cyber physical system. It can be, therefore, assumed that these are leading Industry 4.0 technologies.

4. Summary

The analysis of scientific documents concerning Industry 4.0 with the use of text mining is an interesting method to enrich knowledge in this field. It is an iterative method carried out in stages. It is also time-consuming. It results, first of all, from the need to write the necessary scripts and to correct them when analyzing non-structural texts. On the basis of only 40 articles, technologies appearing in the context of this field could be pre-deduced. It is also possible to attempt to group and compare them. Text mining gives ample opportunities when it comes to analysis and it all depends on the needs and skills of the user. It concerns not only scientists, but also analytics and persons making varied decisions based on the data sources from social and economic environment. Additionally, thanks to the analysis of scientific texts with the use of text mining, it is possible to ask newer and newer questions according to one's own needs on the basis of the obtained results. Therefore, it is so important to develop simple and useful tools and methods of analysis of non-structural text documents. In this article, the assumptions of text mining method used to analyze documents on Industry 4.0 were proposed. Thanks to it, a list of valuable articles on Industry 4.0 was obtained, the PDF files were converted into TXT files, the data was cleared, texts were divided into characteristic parts, the most frequently used words and 2-, 3-words phrases were identified and the correspondence analysis was used to group the selected technologies and articles linked with them.

When analyzing the scientific articles with the use of text mining and stages of proceeding, one should also take into account the limitations imposed by them. They concern, first of all, schemes, tables and drawings found in PDF files that are skipped, but contain many useful information.

The results obtained with the use of definition analysis, schemes and text mining allowed to identify the four currently leading technologies. These are Cloud Computing, IoT (Internet of Things), CPS (Cyber-physical System) and Big Data.

The text mining analysis can be used by specialists of all fields within the Industry 4.0 and can be an interesting addition to the research methods, such as surveys, interviews, observations and literature analysis. In the future, it is necessary to refine the assumptions of this method and analyze a greater number of scientific texts on Industry 4.0.

References

1. Bortolini, M., Ferrari, E., Gamberi, M., Pilati, F., and Faccio, M. (2017). Assembly system design in the Industry 4.0 era: a general framework. *IFAC-PapersOnLine, 50(1),* pp. 5700-5705.
2. Bradley, J., Loucks, J., Macaulay, J., Noronha, A., and Wade, M. (2015). Digital vortex: *How digital disruption is redefining industries*. Global Center for Digital Business Transformation: An IMD and Cisco initiative, p. 6-16. Available online https://www.cisco.com/c/dam/en/us/solutions/collateral/industry-solutions/digital-vortex-report.pdf, 29.10.2019.
3. Fan, W., Wallace, L., Rich, S., and Zhang, Z. (2006). Tapping the power of text mining. *Communications of the ACM, 49(9)*, pp. 76-82.
4. Internet source: *Apple Special Event*. https://www.apple.com/apple-events/. Internet source: Google I/O '19 Recap https://events.google.com/io/recap/, 29.10.2019.
5. *Future Manufacturing Technologies*. https://www.fmtexpo.org/, Industries 4.0 http://www.industries4pointzero.com/conference/page/TC, Smart Industry Expo http://www.smartindustry-expo.com/, 29.10.2019.
6. Moeuf, A., Pellerin, R., Lamouri, S., Tamayo-Giraldo, S., and Barbaray, R. (2018). The industrial management of SMEs in the era of Industry 4.0. *International Journal of Production Research, 56(3)*, pp. 1118-1136.
7. Nakagawa, T., and Uchimoto, K. (2007, June). *A hybrid approach to word segmentation and pos tagging*. Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions, pp. 217-220.
8. Trappey, A.J., Trappey, C.V., Govindarajan, U.H., Chuang, A.C., and Sun, J.J. (2017). A review of essential standards and patent landscapes for the Internet of Things: A key enabler for Industry 4.0. *Advanced Engineering Informatics*, *33*, pp. 208-229.
9. Schumacher, A., Erol, S., and Sihn, W. (2016). *A Maturity Model for Assessing Industry 4.0 Readiness and Maturity of Manufacturing Enterprises*. Procedia CIRP 52, pp. 161-166. doi:10.1016/j.procir.2016.07.040.
10. Saturno M., Pertel V.M., Deschamps F., and Loures, E. (2017). *Proposal of an automation solutions architecture for industry 4.0*. Proceedings of the 24th International Conference on Production Research. Poznań: ICPR.
11. Silge, J., and Robinson, D. (2017). Text mining with R: A tidy approach. *O'Reilly Media, Inc.* Available online https://www.tidytextmining.com/index.html, 10.11.2019.
12. Vijayarani, S., Ilamathi, J., Nithya, and Phil, M. (2015). Preprocessing Techniques for Text Mining – An Overview. *International Journal of Computer Science & Communi-cation Networks, 5(1)*, pp. 7-16.