

Article

# Career in Cloud Computing: Exploratory Analysis of in-Demand Competency Areas and Skill Sets

Ozcan Ozyurt<sup>1</sup>, Fatih Gurcan<sup>2</sup>, Gonca Gokce Menekse Dalveren<sup>3</sup> and Mohammad Derawi<sup>4,\*</sup>

<sup>1</sup> Software Engineering Department, OF Technology Faculty, Karadeniz Technical University, Trabzon 61080, Turkey

<sup>2</sup> Department of Computer Engineering, Faculty of Engineering, Karadeniz Technical University, Trabzon 61080, Turkey

<sup>3</sup> Department of Software Engineering, Faculty of Engineering, Atılım University, Ankara 06830, Turkey

<sup>4</sup> Department of Electronic Systems, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology, 7034 Gjøvik, Norway

\* Correspondence: mohammad.derawi@ntnu.no

**Abstract:** This study aims to investigate up-to-date career opportunities and in-demand competence areas and skill sets for cloud computing (CC), which plays a crucial role in the rapidly developing teleworking environments with the COVID-19 pandemic. In this paper, we conducted a semantic content analysis on 10,161 CC job postings using semi-automated text-mining and probabilistic topic-modeling procedures to discover the competency areas and skill sets as semantic topics. Our findings revealed 22 competency areas and 46 skills, which reflect the interdisciplinary background of CC jobs. The top five competency areas for CC were identified as “Engineering”, “Development”, “Security”, “Architecture” and “Management”. Besides, the top three skills emerged as “Communication Skills”, “DevOps Tools” and “Software Development”. Considering the findings, a competency-skill map was created that illustrates the correlations between CC competency areas and their related skills. Although there are many studies on CC, the competency areas and skill sets required to deal with cloud computing have not yet been empirically studied. Our findings can contribute to CC candidates and professionals, IT organizations, and academic institutions in understanding, evaluating, and developing the competencies and skills needed in the CC industry.

**Keywords:** cloud computing; skill requirements; competency areas; topic modeling; job-posting analysis; text mining



**Citation:** Ozyurt, O.; Gurcan, F.; Dalveren, G.G.M.; Derawi, M. Career in Cloud Computing: Exploratory Analysis of in-Demand Competency Areas and Skill Sets. *Appl. Sci.* **2022**, *12*, 9787. <https://doi.org/10.3390/app12199787>

Academic Editor: Pentti Nieminen

Received: 18 August 2022

Accepted: 27 September 2022

Published: 28 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The COVID-19 pandemic and the resulting economic recession have accelerated the transition of many companies, institutions, and governments to cloud computing (CC) because cloud services and applications provide the flexibility and scalability that can help them to manage such crises more effectively. With the emergence of the COVID-19 pandemic, many companies and institutions have embraced teleworking more than before, leveraging CC-based services and applications to reduce the risk of virus transmission and manage the current crisis [1]. In this pandemic era that restricts working and social-life conditions, CC has become one of the most prominent information technologies thanks to the solutions and modernizations it provides in the delivery of computing resources and services [2,3]. Therefore, cloud computing (CC), which is envisaged as the leading paradigm of resource and service delivery, is also considered the next generation of computing [4–6]. CC offers significant flexibility in pricing computing services by allowing organizations to only pay for the resources and services they use [7,8]. Therefore, CC, which is accepted as next-generation computing, has become the main paradigm of resource and service delivery [4–6]. In addition, CC provides many innovations such as advanced institutional visibility, easier collaboration, rapid development, consistency across platforms, enhanced security, lower risk,

and reduced cost [5,6,9]. Thanks to these innovations, CC is accepted as the fifth service after four basic services, including water, electricity, gas, and telephone [6,10].

In line with these modernizations, the adoption of CC technologies has been increasing rapidly in recent years [6,11]. It is estimated that end-user spending on public cloud services globally will increase by 23.1%, from USD 332.3 billion in 2021 to USD 397.5 billion in 2022 [12]. The COVID-19 pandemic has also had a major impact on this growth, and the demand for critical CC resources and services has increased exponentially during the pandemic [13]. In another view, cloud platforms and ecosystems will serve as a launching pad for a future explosion of digital innovation in the next five years [6]. In this context, the CC industry will become even more important for IT professionals in the near future [14]. These dizzying developments in the CC industry will further increase the demand for a skilled workforce [15]. The CC industry has a dynamic and entrepreneurial working environment in which capable human resources are used effectively [9]. Therefore, today's CC job market offers significant employment and career opportunities for IT professionals [9]. One of the dynamic information sources about competencies, knowledge, and skills demanded by the CC job market is online employment platforms [15,16]. These platforms are also important communication tools between employees and employers. A large number of CC job postings are posted on these platforms every day. These job postings can be considered an important indicator that reveals emerging needs and trends in today's CC job market [17–19].

In this respect, job postings, which is used for employment in this field, can be seen as an indicator of CC industry needs and trends [15,16]. Therefore, analysis of job postings is important not only to identify the competencies demanded for the workforce, but also to identify current trends in the CC industry [16]. With regard to the analysis of job postings, numerous studies have been conducted on different sub-contexts of IT jobs using various text-mining techniques [3,19,20]. In analyzing job postings, these studies mostly aimed to reveal the competencies and skills required for subfields of IT such as computer science, big data, software engineering, computer programming, and big-data software engineering [15–17,21–23]. On the other hand, it is possible to mention a few articles on the identification of CC competencies and the inclusion of these competencies in the curricula of IT departments, such as computer science and software engineering [8,15,24–26]. These studies have highlighted the gap in the skilled workforce needed in the CC industry in general. In this regard, these valuable studies and efforts to reveal and meet the competent-workforce gap in the CC industry constitute the motivation and background of our study. Although many studies have been conducted on CC in different contexts, the competencies and skill sets required for cloud computing have not yet been empirically studied.

Considering this background, this study aimed to identify the knowledge domains, skill sets, and competencies demanded by the CC industry. In this regard, a semantic content analysis was performed on CC job postings using text-mining procedures and a probabilistic topic-modeling approach based on latent Dirichlet allocation (LDA). As a result, this analysis revealed the 22 main competency areas and 46 trending topics (skills) essential for CC jobs. Afterwards, a competency taxonomy for CC was developed by calculating the distribution of discovered topics (skills) according to competency areas. From this point of view, it is expected that identifying and understanding the up-to-date knowledge domains and skill sets demanded by the CC industry will offer valuable contributions to meet the need for a qualified workforce in this field.

## 2. Materials and Methods

Consistent with the purpose of the study, the research methodology was designed based on topic-modeling processes that reveal the main dynamics, themes, and trends of CC's vibrant job market as topics discovered with quantitative procedures. In this study, a semi-automatic methodology was used for semantic-content analysis based on topic modeling of online CC job postings. In this context, this methodology consisted of three

main phases: data collection, data preprocessing, data analysis, and interpretation. Each phase of the methodology is explained in more detail in the following sub-headings.

### 2.1. Data Collection and Preprocessing

With the aim of creating the experimental dataset of the study, initially we searched for a reasonable data source (online-employment site) from which we could obtain job postings effectively. Among these identified sources, Indeed.com, an online-employment site that offers extensive search-and-filtering options, was chosen as the data source for this study. Indeed.com is the world's #1 job site with over 250 million unique visitors every month [27]. In this context, in order to include the most appropriate job postings for the dataset, all postings with the term "cloud" in the job title were considered CC jobs [28]. Sample job postings that serve as a data source for the construction of our empirical dataset can be viewed on the Indeed website [28] using the search query we created [28]. In this way, 10161 CC job postings published in a quarterly period from 15 November 2021 to 15 February 2022 were extracted using an API developed by Indeed.com. The extracted job postings were saved in an Excel file (one job posting per line) and thus the experimental data set of the study were created. In our dataset, each job posting consists of two sub-sections (represented in two columns) containing the job title and job description.

Following the creation of the data set, data-preprocessing steps, which is one of the main tasks of text mining, were implemented on the textual data. Specifically, common preprocessing steps necessary for text analysis, including tokenization (splitting the full text into singular words), stop-word removal (a, the, of, for, etc.), lowercase conversion, elimination of missing and misspelled words, and lemmatization processes were implemented. After the preprocessing steps, the document-term matrix (DTM) was constructed with the remaining words in order to achieve the intended numerical analysis. Thanks to the preprocessing steps mentioned above, the lower dimensional DTM was able to characterize the dataset. A DTM is a numerical matrix that reveals the frequency of the terms that occur in a textual corpus.

### 2.2. Data Analysis and Interpretation

In the first stage of the empirical analysis, job titles were categorized according to their respective competency areas. The job title of each CC job is an important indicator that specifies the competency area of that job. From this point of view, considering the words in the job titles, each title was assigned to a competency area related to CC. For example, job titles such as Cloud Architect, AWS Cloud Architect, Azure Cloud Architect, or Cloud Architecture Expert, which contain derivatives of the word "architecture", are included in the competency area of architecture. In the same way, the words contained in other CC job titles were identified and all job titles were assigned to their respective competency area [15,16]. In this way, 22 fundamental competency areas for CC jobs were identified.

In the second stage of the analysis, semantic text-mining procedures based on probabilistic topic modeling were performed on the textual corpus of the CC job postings. In semantic text mining, topic modeling is a generative probabilistic approach used for analysis of the semantic structure of a textual corpus [29]. Text documents contain latent semantic patterns called "topics". A text document may contain multiple topics with different ratios. The identification of these topics and their ratios constitutes the basis of topic models. Latent Dirichlet allocation (LDA) [29,30] is a common topic-modeling algorithm used in semantic text mining. LDA consists of a set of matrix operations performed on the DTM in order to identify semantic correlations between the words. LDA is an unsupervised learning model that can extract latent topics from a textual corpus without any training or annotation [31,32]. Therefore, the LDA algorithm was employed for topic-modeling analysis in this study.

With the aim of fitting and implementing the LDA model to the corpus of CC jobs, the `tmtoolkit` package [33], an extensive toolkit developed in Python for topic modeling, was employed. The prior parameters of  $\alpha$ , which sets the topic dissemination in the documents,

and  $\beta$ , which sets the word dissemination in the topics, were selected as  $\alpha = 0.1$  and  $\beta = 0.01$ , considering related topic-modeling studies [31,34,35]. Subsequently, with the intention of choosing the optimal number of topics (K), the LDA model was implemented with different K values ranging from 10 to 60. For each K value changing in this range, a coherence score for each fitted model was calculated [33,36]. Considering the coherence scores calculated for each K in this range, we observed that as the number of topics increased, the consistency score generally decreased inversely. In these calculated scores, the threshold values were evaluated for each experiment. A maximum coherence score, which denotes the finest semantic consistency of the discovered topics, was achieved for the topic number of  $K = 46$  [36]. The top 15 descriptive keywords with the highest frequency were identified for each of the 46 topics. Afterwards, the semantic consistency of the topics was re-evaluated by the two field experts and the topic name of each topic was identified and assigned considering the descriptive keywords of the topics. In addition, the distribution rates of the discovered topics per document, the distribution of words per topic, and the distribution rates of the topics in the entire corpus were calculated [15,30,33]. Consequently, these 46 topics discovered by LDA were used in all following analyses.

### 3. Results

In this section, the results obtained from the analysis of CC job postings are given. Competency areas and related job titles were initially identified from job postings. A total of 22 competency areas were identified from job postings considering frequency of related job titles. These competency areas and the first five job titles related to them, as well as their number and percentages, are given in Table 1.

Table 1 indicates that a total of 22 competency areas were obtained from CC job postings. The first five of these areas emerged as “Engineering”, “Development”, “Security”, “Architecture”, and “Management”. The ratio of the first five competencies in the total jobs was calculated as 51.48%. In other words, more than half of the CC jobs referred to the first five competencies. In addition, each line shows the first five job titles required for the relevant competency in order. In Figure 1, the top 30 job titles obtained from all jobs are given in order of percentile volume. A total of 875 different job titles were identified, and the first 30 titles accounted for 25.14% of the total titles. The first three titles were identified as “Cloud Engineer” ( $n = 705$ ;  $f = 6.94\%$ ), “Cloud Architect” ( $n = 265$ ;  $f = 2.61\%$ ), and “Cloud Security Engineer” ( $n = 159$ ;  $f = 1.56\%$ ), respectively.

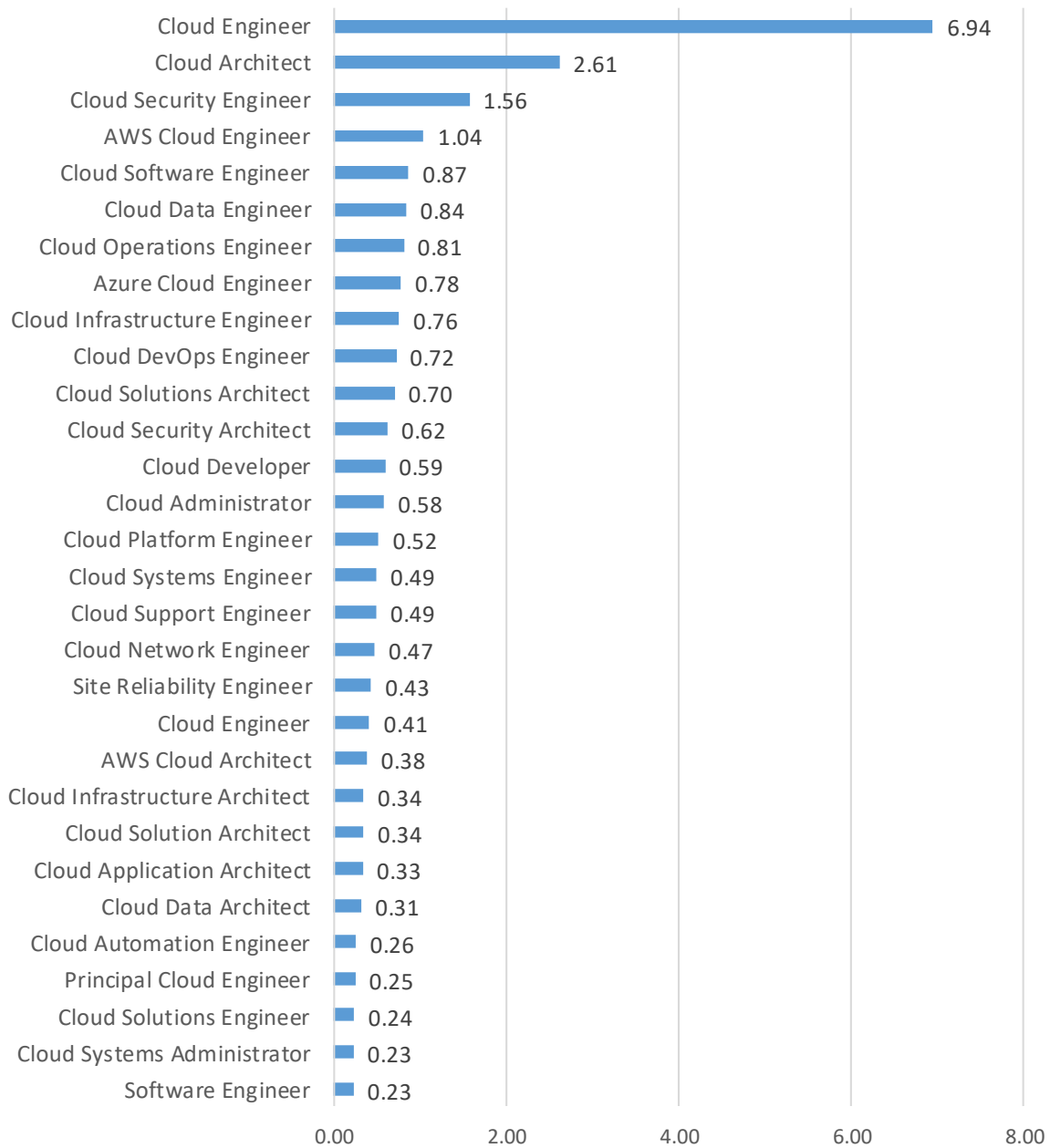
In this part of the results section, the results obtained from the topic-modeling analysis based on LDA are given. As a result of this topic-modeling analysis, a total of 46 topics (skills) were discovered. These topics are presented in Table 2 according to the order of percentage. These 46 topics discovered by topic-modeling analysis can be considered skills in this field. In this respect, “topic” and “skill” are used synonymously throughout the text.

Table 2 indicates that the three most voluminous topics were “Communication Skills” (5.02%), “DevOps Tools” (4.89%), and “Software Development” (4.48%), whereas the least voluminous topics were “Infrastructure Solutions” (0.85%), “Global Solutions” (0.76%), and “Delivery Solutions” (0.64%). The topic names, their descriptive keywords, and the percentages of the topics are presented thoroughly in Appendix A.

In addition, the relationship between all topics (skills) and the first 10 competency areas is presented in Table 3 as a percentage. “Communication Skills”, which was the most voluminous topic, had a volume of 5.02% among all topics. The percentage of this topic in the top 10 competencies can also be seen in the first line of Table 3. For example, “Communication Skills” emerged as the most in-demand skill in the “Engineering” competency area, with a rate of 0.8%. This was followed by “Development”, with 0.62%. In other words, “Communication Skills” was most in demand from the “Engineering” competency area. The area of “Development” was in second place among the areas that demanded this skill. The competency area where “Communication Skills” was least desired or needed was “DevOps” (within the top 10 areas), with 0.15%.

**Table 1.** Competency areas, related job titles, numbers, and percentages.

Competency Area	The Most In-Demand Job Titles	n	%
Engineering	Cloud Engineer, AWS Cloud Engineer, Azure Cloud Engineer, Google Cloud Engineer, Principal Cloud Engineer	1662	16.36
Development	Software Engineer, Cloud Software Engineer, Cloud Developer, Cloud Application Architect, Software Engineering	1397	13.75
Security	Cloud Security Engineer, Cloud Security Architect, Cloud Security Analyst, Cloud Security Operations Engineer, Cloud Security Specialist	895	8.81
Architecture	Cloud Architect, AWS Cloud Architect, Azure Cloud Architect, Principal Cloud Architect, Google Cloud Architect	686	6.75
Management	Product Manager, Project Manager, Engineering Manager, Program Manager, Cloud Engineering Manager	590	5.81
Data	Cloud Data Engineer, Cloud Data Architect, Cloud Database Administrator, Cloud Storage Engineer, Virtual Cloud Data Engineer	561	5.52
Solution	Cloud Solution Architect, Cloud Solution Engineer, Solution Architect, Azure Cloud Solution Architect, AWS Cloud Solution Architect	468	4.61
Infrastructure	Cloud Infrastructure Engineer, Cloud Infrastructure Architect, Infrastructure Engineer, IT Infrastructure Design, Cloud Infrastructure Services	468	4.61
DevOps	Cloud DevOps Engineer, DevOps Cloud Engineer, DevOps Engineer, DevOps/Cloud/Infra Consultant, Cloud DevOps Infrastructure Engineer	387	3.81
Services	Cloud Services Engineer, Apple Cloud Services, Cloud Services Solution Engineer, Google Cloud Professional Services, Cloud Engineer for Cloud Services	329	3.24
Sales	Cloud Sales Specialist, Sales Engineer, Cloud Sales Representative, Sales Development Representative, Enterprise Cloud Sales Representative	306	3.01
Analytics	Cloud Data Analyst, Business Analyst, Data Analyst, Cloud Business Analyst, Cloud Cost Optimization Analyst	288	2.83
Administration	Cloud Administrator, Cloud System Administrator, System Administrator, Oracle Cloud Administrator, Cloud Operations Administrator	275	2.71
Support	Cloud Support Engineer, Technical Support Engineer, Cloud Support Specialist, Cloud Operation Support Specialist, Cloud Application Support Engineer	274	2.70
Consultancy	Cloud Consultant, Oracle Cloud HCM Consultant, Oracle Cloud Supply Chain Consultant, Oracle ERP Cloud Consultant, Associate Cloud Consultant	271	2.67
Operations	Cloud Operations Engineer, Cloud Ops Engineer, Cloud Operations Manager, Director Cloud Operations, Azure Cloud Operations Engineer	262	2.58
Marketing	Product Marketing Manager, Cloud Marketing Engineer, Cloud Marketing Solutions Architect, Cloud Marketing Specialist, Cloud Marketing Consultant	229	2.25
System	Cloud System Engineer, System Engineer, Cloud System Architect, Cloud System Management Specialist, AWS Cloud System Engineer	198	1.95
Finance	Oracle Cloud Finance Consultant, Finance Manager, Oracle ERP Cloud Financial Consultant, Financial Analyst, Cloud FinOps Analyst	172	1.69
Platform	Cloud Platform Engineer, Cloud Platform Architect, Google Cloud Platform Engineer, Cloud Core Platform Engineer, Cloud Ops Platform Engineer	170	1.67
Network	Cloud Network Engineer, Network Cloud Engineer, Cloud Network Architect, Cloud Networking Engineer, Azure Cloud Network Engineer	149	1.47
Automation	Cloud Automation Engineer, Cloud Automation Software Engineer, Cloud Automation Architect, Cloud Delivery Automation Engineer, Cloud Automation Expert	124	1.22



**Figure 1.** Top 30 job titles and percentages.

Table 4 presents the relationship between importance of the topics (skills) and the competency areas from a different perspective. The order of importance of the top 10 topics (skills) required for each of the top 10 competency areas is given in this table. Table 4 can be read based on a row or a column. Each line indicates the order of that skill in the related competency areas. Each column indicates the order of the top 10 topics in that competency area.

**Table 2.** Topics (skills) discovered by LDA.

ID	Topic Name	%	ID	Topic Name	%
1	Communication Skills	5.02	24	Health Care	1.89
2	DevOps Tools	4.89	25	Customer Services	1.79
3	Software Development	4.48	26	AWS	1.67
4	Architectural Solutions	4.18	27	Network Infrastructure	1.66
5	Service Support	4.13	28	Consulting Services	1.61
6	Production Management	3.84	29	Financial Solutions	1.50
7	Bachelor's Degree	3.68	30	IBM Cloud	1.46
8	Project Management	3.51	31	Digital Marketing	1.32
9	Strategy Development	3.43	32	Client Solutions	1.24
10	Problem Solving	3.40	33	Cost Analysis	1.23
11	Service Models	3.34	34	Remote Development	1.20
12	Cybersecurity	2.85	35	Data Center Solutions	1.20
13	Public Services	2.79	36	Essential Skills	1.14
14	Technical Knowledge	2.69	37	Network Security	1.02
15	Information Services	2.42	38	Information Processing	1.01
16	Government Services	2.30	39	Training Certification	0.99
17	Sales Solutions	2.28	40	Service Requirements	0.99
18	Identity Protection	2.28	41	Database	0.92
19	Teamwork	2.25	42	VMware Services	0.90
20	Big Data Analytics	2.16	43	Oracle Services	0.85
21	Azure Services	2.11	44	Infrastructure Solutions	0.78
22	Google Cloud	2.10	45	Global Solutions	0.76
23	Oracle ERP	2.09	46	Delivery Solutions	0.64

**Table 3.** Percentages of skills in the top 10 competencies.

Topic Name/Competency Area	Engineering	Development	Security	Architecture	Management	Data	Solution	Infrastructure	DevOps	Services
Communication Skills	0.80	0.62	0.44	0.30	0.34	0.23	0.25	0.21	0.15	0.18
DevOps Tools	1.47	0.59	0.34	0.34	0.06	0.13	0.12	0.27	0.64	0.08
Software Development	0.78	1.97	0.18	0.24	0.09	0.20	0.11	0.09	0.19	0.11
Architectural Solutions	0.68	0.38	0.36	0.89	0.11	0.23	0.41	0.23	0.11	0.10
Service Support	0.92	0.22	0.23	0.15	0.09	0.16	0.09	0.27	0.20	0.12
Production Management	0.32	0.36	0.29	0.23	0.68	0.19	0.29	0.14	0.06	0.20
Bachelor's Degree	0.83	0.68	0.30	0.26	0.14	0.16	0.18	0.14	0.14	0.12
Project Management	0.36	0.39	0.24	0.20	0.56	0.16	0.15	0.17	0.09	0.24



Table 3. Cont.

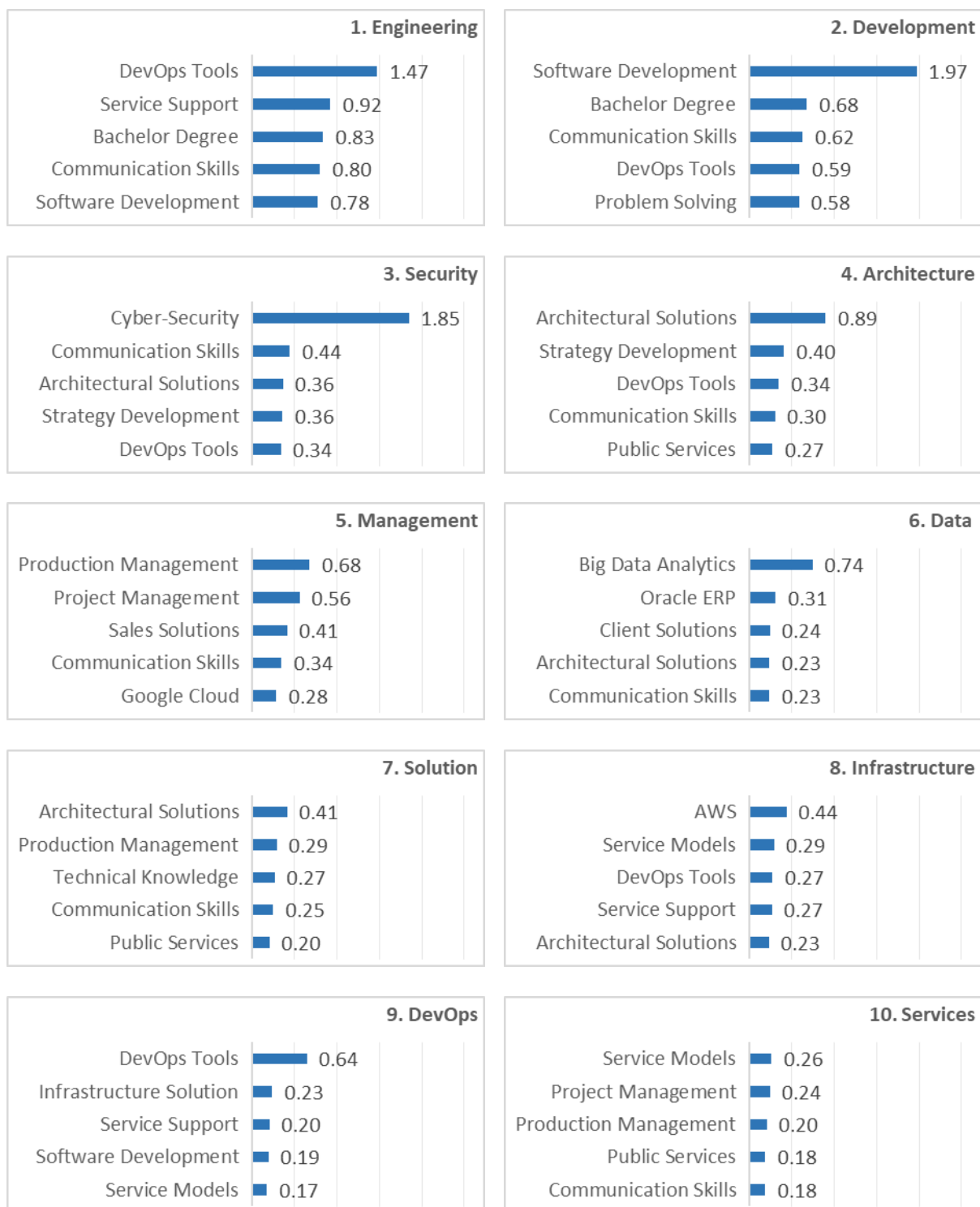
Topic Name/Competency Area	Engineering	Development	Security	Architecture	Management	Data	Solution	Infrastructure	DevOps	Services
Strategy Development	0.60	0.33	0.36	0.40	0.16	0.19	0.18	0.16	0.11	0.13
Problem Solving	0.59	0.58	0.26	0.23	0.19	0.18	0.14	0.17	0.15	0.11
Service Models	0.64	0.55	0.30	0.13	0.18	0.14	0.08	0.29	0.17	0.26
Cybersecurity	0.25	0.08	1.85	0.08	0.06	0.03	0.03	0.07	0.06	0.03
Public Services	0.41	0.26	0.21	0.27	0.21	0.12	0.20	0.11	0.08	0.18
Technical Knowledge	0.38	0.34	0.21	0.21	0.09	0.15	0.27	0.10	0.07	0.08
Information Services	0.41	0.34	0.23	0.18	0.11	0.11	0.07	0.10	0.06	0.07
Government Services	0.60	0.27	0.17	0.22	0.04	0.08	0.10	0.03	0.09	0.08
Sales Solutions	0.05	0.12	0.05	0.04	0.41	0.03	0.20	0.02	0.01	0.15
Identity Protection	0.45	0.32	0.23	0.15	0.13	0.12	0.09	0.09	0.07	0.07
Teamwork	0.38	0.41	0.30	0.10	0.12	0.12	0.07	0.11	0.09	0.05
Big Data Analytics	0.30	0.28	0.04	0.12	0.05	0.74	0.07	0.04	0.07	0.02
Azure Services	0.65	0.13	0.15	0.17	0.05	0.06	0.12	0.11	0.07	0.05
Google Cloud	0.16	0.18	0.12	0.07	0.28	0.09	0.15	0.07	0.02	0.13
Oracle ERP	0.04	0.13	0.03	0.07	0.17	0.31	0.04	0.02	0.01	0.04
Health Care	0.38	0.28	0.14	0.13	0.08	0.11	0.06	0.11	0.07	0.03
Customer Services	0.23	0.09	0.11	0.05	0.07	0.14	0.07	0.03	0.07	0.04
AWS	0.07	0.42	0.06	0.16	0.11	0.02	0.05	0.44	0.02	0.03
Network Infrastructure	0.32	0.12	0.10	0.09	0.02	0.06	0.05	0.18	0.05	0.03
Consulting Services	0.18	0.11	0.19	0.11	0.13	0.09	0.11	0.06	0.10	0.05
Financial Solutions	0.17	0.54	0.11	0.09	0.06	0.15	0.06	0.03	0.04	0.03
IBM Cloud	0.06	0.29	0.17	0.13	0.09	0.12	0.08	0.09	0.03	0.03
Digital Marketing	0.08	0.35	0.04	0.07	0.09	0.03	0.04	0.01	0.02	0.04
Client Solutions	0.22	0.14	0.06	0.07	0.05	0.24	0.02	0.09	0.04	0.02
Cost Analysis	0.10	0.07	0.08	0.04	0.13	0.06	0.04	0.06	0.02	0.03
Remote Development	0.27	0.16	0.11	0.12	0.04	0.07	0.03	0.04	0.07	0.02
Data Center Solutions	0.19	0.30	0.09	0.05	0.10	0.05	0.10	0.03	0.05	0.01
Essential Skills	0.21	0.14	0.08	0.06	0.06	0.05	0.04	0.04	0.04	0.04
Network Security	0.11	0.15	0.20	0.02	0.05	0.01	0.06	0.02	0.02	0.02
Information Processing	0.18	0.07	0.05	0.07	0.09	0.08	0.03	0.04	0.02	0.03
Training Certification	0.46	0.11	0.07	0.09	0.01	0.01	0.02	0.03	0.04	0.01
Service Requirements	0.24	0.18	0.08	0.07	0.02	0.05	0.05	0.04	0.03	0.02
Database	0.21	0.19	0.03	0.08	0.02	0.13	0.03	0.04	0.04	0.02
VMware Services	0.09	0.10	0.05	0.05	0.09	0.03	0.05	0.09	0.02	0.05
Oracle Services	0.05	0.08	0.01	0.03	0.05	0.04	0.05	0.02	0.01	0.01
Infrastructure Solutions	0.19	0.10	0.02	0.03	0.01	0.01	0.04	0.06	0.23	0.01
Global Solutions	0.08	0.10	0.04	0.03	0.08	0.03	0.09	0.02	0.05	0.04
Delivery Solutions	0.20	0.13	0.04	0.07	0.03	0.02	0.01	0.04	0.02	0.02



**Table 4.** Ranking of the top 10 skills for each of the top 10 competency areas.

Skill/Competency Area	Engineering	Development	Security	Architecture	Management	Data	Solution	Infrastructure	DevOps	Services
Communication Skills	4	3	2	4	4	5	4	6	6	5
Strategy Development	9		4	2	10	7	8	10	9	8
Architectural Solutions	6		3	1		4	1	5	10	
Bachelor's Degree	3	2	8	6			7		8	9
Problem Solving		5	10	9	7	9		8	7	
Service Models	8	6	7		8			2	5	1
DevOps Tools	1	4	5	3				3	1	
Production Management			9	8	1	8	2			3
Project Management		10			2		9	9		2
Service Support	2					10		4	3	10
Software Development	5	1		7		6			4	
Public Services				5	6		5			4
Google Cloud					5		10			7
Sales Solutions					3		6			6
AWS		8						1		
Government Services	10			10						
Oracle ERP					9	2				
Teamwork		9	6							
Azure Services	7									
Big Data Analytics						1				
Client Solutions						3				
Cybersecurity			1							
Financial Solutions		7								
Infrastructure Solutions									2	
Network Infrastructure								7		
Technical Knowledge							3			

For example, in line 1 in the table, “Communication Skills” was a necessary topic for the top 10 competency areas and its rank for each area ranged from 2 to 6. Similarly, the topic of “Big Data Analytics” was the only skill in the “Data” competency area and was in first place. The top 10 topics of each competency area are given as column-based reading. Namely, for column 1, the top 10 skills of the “Engineering” competency area are presented in that column, with the first of them being “DevOps Tools”. On the other hand, the second and third are “Bachelor’s Degree” and “Service Support”, respectively. Finally, the top five topics (skills) of the top 10 competency areas are demonstrated with their percentages in Figure 2. This chart offers a graphical illustration of the top five skills required for each of the top 10 areas. For example, as seen in the chart, the first skill of the “Engineering” competency area was “DevOps Tools”, whereas this skill was ranked fourth for the “Development” area.



**Figure 2.** Percentage charts of the top five skills in the top 10 competency areas.

#### 4. Discussion

In this study, a semantic content analysis was conducted on CC job postings using text-mining and probabilistic topic-modeling processes. As a result of this analysis, the 22 main competency areas and 46 skills necessary for CC jobs were revealed. The top five competency areas were “Engineering”, “Development”, “Security”, “Architecture”, and “Management.” Brievold and Crnkovic (2014) listed the fields of knowledge in cloud computing as (a) cloud fundamentals, (b) architecting for the cloud, (c) design for resilience,

(d) cloud security and compliance, and (e) migrating to the cloud [37]. In their study, “Architecture” and “Security” became prominent in addition to the basic concepts of the CC. The fact that “Security” and “Architecture” competency areas were at the top in the current study highlights the significance of these areas for CC. In addition to technical areas such as “Engineering”, “Development”, and “Architecture”, the dominance of skills related to administrative processes and paradigms emphasizes the importance of the “Management” competency area. The 22 competency areas required for CC professions and their specific skills will be the key qualifications demanded by many technology-related careers in the age of Industry 4.0, as highlighted in the literature [25].

A total of 875 different job titles, constituting the 22 competency areas required in CC jobs, shows that the qualifications and skills needed for the CC industry cover a wide spectrum. Although a limited number of these job titles have been revealed in some studies [25], this current study revealed a wide spectrum of diversity of CC job titles. The emergence of a large number of different job titles confirms the magnitude of CC’s workforce requirements and its dominance in today’s IT industries [7]. The development and dominance of cloud technologies enable cloud engineering to be a lively competency area that is constantly improving and maturing for the CC industry, as specified in the literature and this current study [38]. In addition, CC requires a strong infrastructure and competency in service-oriented network architecture to achieve large-scale and complex computing services. Cloud computing is a comprehensive collection of many different computing paradigms and technologies, encompassing grid computing, virtualization, autonomous computing, service-oriented architecture (SOA), peer-to-peer (P2P) computing, and ubiquitous computing needed in different service platforms [39]. In this context, the implications obtained from our analysis also indicates the important role of cloud architecture in the CC industry [14].

The results of our topic-modeling analysis revealed 46 topics that illustrate the domain specific knowledge and skills that are highly in demand for the CC industry. The findings of this analysis indicate that expertise in the CC industry requires a wide-ranging spectrum of knowledge domains, skill sets, and abilities. In line with these findings, a competency matrix with a taxonomic perspective was created to organize these competency areas and their related skills. These discovered competencies and skills reveal that CC expertise requires the combined use of a comprehensive collection of both technical and soft-skills usage [6,15]. Considering these skills, “DevOps Tools”, “Software Development”, and “Architectural Solutions” are the leading technical skills specific to the CC field. Among our findings, “Communication Skills” ranked first, “Bachelor’s Degree” ranked seventh, and “Problem Solving” ranked tenth as non-technical skills. In today’s competitive labor market, the need for IT professionals to have technical and social-communication skills has also been discussed in different aspects in some studies [40]. In addition, some previous studies have drawn attention to the importance of problem-solving skills as well as communication skills in IT education [15,40–42].

## 5. Conclusions

In this study, a semantic content analysis was performed on CC job postings using text-mining and topic-modeling procedures in order to reveal the competency areas, knowledge domains, and skills sets demanded by the CC industry. This study found that (1) the discovered competency areas and skills also reveal the emerging trends and demands in the CC industry, as well as the required qualifications for CC professionals; (2) CC expertise requires a wide spectrum of knowledge, skills, and abilities with an interdisciplinary background; and (3) as leading actors, CC professionals can undertake different roles (22 competency areas and 875 different job titles) that require comprehensive skill sets that combine many technical and soft skills (46 different skills).

Since there is no experimental study in this context in the literature, this study is expected to make significant contributions to CC communities. The findings of this study can provide valuable insights into the understanding of the main characteristics and require-

ments of CC jobs. These findings may offer meaningful implications for CC stakeholders from different aspects. At the institutional level, the findings may help cloud companies to identify qualified CC professionals, and they may help academic institutions meet the need for a qualified CC workforce. At the individual level, the findings may be helpful for CC professionals in measuring and updating their own competencies, for instructors in educating CC candidates in line with emerging demands, and for students in scheduling their career paths. In addition, the methodology of this study can be used in future research to reveal the needs and trends of different IT industries.

**Author Contributions:** Conceptualization, F.G. and O.O.; methodology, G.G.M.D. and F.G.; software, G.G.M.D.; validation, G.G.M.D. and O.O.; investigation, G.G.M.D. and M.D.; resources, G.G.M.D. and M.D.; data curation, F.G. and O.O.; writing—original draft preparation, G.G.M.D. and M.D.; writing—review and editing, M.D.; visualization, F.G. and O.O.; supervision, M.D.; project administration, M.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data are available and can be shared upon request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Discovered topics and their keywords and rates.

Topic Name	Topic Keywords	%
Communication Skills	skill ability strong communication excellent write knowledge environment demonstrate understand verbal able problem level effectively	5.02
DevOps Tools	DevOps infrastructure tool engineer AWS automation deployment terraform Kubernetes CI/CD environment code python application container	4.89
Software Development	software development design application test developer develop code product Java engineer service web agile testing	4.48
Architectural Solutions	architecture architect solution design application technical infrastructure enterprise platform AWS implementation security service development migration	4.18
Service Support	support service infrastructure environment operation management application maintain server issue administration monitor provide performance configuration	4.13
Production Management	product drive management partner strategy manager success customer deliver leadership need organization market key technical	3.84
Bachelor's Degree	engineering computer degree science engineer technical software relate solution design development field qualification information equivalent	3.68
Project Management	project management manage program manager process ensure delivery service technical support development requirement quality activity	3.51
Strategy Development	develop strategy design strategic development provide practice process organization implement leadership senior standard support ensure	3.43
Problem Solving	learn solve problem challenge deliver create idea innovation build global innovative talent passion part industry	3.40
Service Models	platform engineering infrastructure service engineer build scale software reliability distribute operation improve building product operate	3.34
Cybersecurity	security risk information control compliance management threat cyber secure vulnerability policy cybersecurity access data network	2.85
Public Services	service solution public transformation azure delivery manage deliver strategy provider digital development support partner hybrid	2.79
Technical Knowledge	technical knowledge solution requirement expert provide expertise project understand matter subject design implementation consult need	2.69

Table A1. Cont.

Topic Name	Topic Keywords	%
Information Services	information service provide require application requirement education resume process need change staff document relate qualification	2.42
Government Services	government service mission support AWS customer provide require security solution environment certification DOD secret intelligence	2.30
Sales Solutions	sale customer solution partner executive account drive marketing sell relationship revenue growth market software develop	2.28
Identity Protection	protect qualify information identity receive application characteristic reasonable require marital applicable orientation expression process local	2.28
Teamwork	team environment work interact member need support remote product background provide community create success build	2.25
Big Data Analytics	data analytic big platform SQL database pipeline tool warehouse ETL model spark python process develop	2.16
Azure Services	Azure Microsoft service directory active engineer solution server security management window infrastructure environment powershell virtual	2.11
Google Cloud	Google customer partner need product qualification platform technical solution infrastructure build organization legal proud policy	2.10
Oracle ERP	oracle client ERP implementation design functional process application management financial project requirement HCM consultant solution	2.09
Health Care	health medical care healthcare vision program support provide patient assistance package competitive coverage account generous	1.89
Customer Services	customer support service AWS technical issue engineer Amazon provide troubleshooting web application critical understand training	1.79
AWS	AWS customer service Amazon professional technical member project balance application development develop learn data need	1.67
Network Infrastructure	network infrastructure knowledge firewall networking security CISCO engineer protocol storage DN load virtualization design VPN	1.66
Consulting Services	service provide consult professional client build support consultant industry deliver need environment engagement success member	1.61
Financial Solutions	financial bank banking basis consumer global solution investment knowledge practice chase applicable management innovative protect	1.50
IBM Cloud	IBM client program IBMer eligible expertise resource support build professional organization hybrid technical learn service	1.46
Digital Marketing	marketing digital content commerce platform data campaign media email create brand studio developer adobe market	1.32
Client Solutions	client center solution development professional delivery design support engineering azure AWS requirement service qualification build	1.24
Cost Analysis	financial cost finance analysis report data process management provide support optimization model analyst performance tool	1.23
Remote Development	remote multiple development communication write project level pattern mantech client current maven verbal ARA senior	1.20
Data Center Solutions	data center global solution software service customer industry platform orientation fortune protect information identity qualification	1.20
Essential Skills	require perform essential skill function able environment physical ability equipment service computer requirement responsibility need	1.14
Network Security	network protect mission security information product cybersecurity need provide challenge political qualify medical description customer	1.02
Information Processing	information process access provide retrieval investigation client statement need require part function discriminate action essential	1.01
Training-Certification	tool certification training course technical engineer build resource develop need challenge access environment security engineering	0.99
Service Requirements	requirement service customer require information remote description qualification policy engineering ID resume mission date seek	0.99
Database	database SQL data MySQL MongoDB customer server relational storage PostgreSQL high Kafka performance Postgre NoSQL	0.92

Table A1. Cont.

Topic Name	Topic Keywords	%
VMware Services	VMware service discrimination provide software customer requirement environment protect solution social union consistent orientation need	0.90
Oracle Services	Oracle need skill qualify service product create protect ability provide PWC receive development record application	0.85
Infrastructure Solution	infrastructure solution technical client application challenge software ensure support create platform innovative tech tool community	0.78
Global Solution	solution global innovation customer software development service background industry build application provide worldwide ensure create	0.76
Delivery Solution	deliver solution customer build express platform square need gold remote security embrace mission multiple member	0.64

## References

- Alashhab, Z.R.; Anbar, M.; Singh, M.M.; Leau, Y.B.; Al-Sai, Z.A.; Alhayja'a, S.A. Impact of Coronavirus Pandemic Crisis on Technologies and Cloud Computing Applications. *J. Electron. Sci. Technol.* **2021**, *19*, 100059. [CrossRef]
- Nadiv, R. Home, Work or Both? The Role of Paradox Mindset in a Remote Work Environment during the COVID-19 Pandemic. *Int. J. Manpow.* **2021**, *43*, 1182–1203. [CrossRef]
- Karakolis, E.; Kapsalis, P.; Skalidakis, S.; Kontzinos, C.; Kokkinakos, P.; Markaki, O.; Askounis, D. Bridging the Gap between Technological Education and Job Market Requirements through Data Analytics and Decision Support Services. *Appl. Sci.* **2022**, *12*, 7139. [CrossRef]
- Qasem, Y.A.M.; Abdullah, R.; Jusoh, Y.Y.; Atan, R.; Asadi, S. Cloud Computing Adoption in Higher Education Institutions: A Systematic Review. *IEEE Access* **2019**, *7*, 63722–63744. [CrossRef]
- Senyo, P.K.; Addae, E.; Boateng, R. Cloud Computing Research: A Review of Research Themes, Frameworks, Methods and Future Research Directions. *Int. J. Inf. Manag.* **2018**, *38*, 128–139. [CrossRef]
- Varghese, B.; Buyya, R. Next Generation Cloud Computing: New Trends and Research Directions. *Futur. Gener. Comput. Syst.* **2018**, *79*, 849–861. [CrossRef]
- Assunção, M.D.; Calheiros, R.N.; Bianchi, S.; Netto, M.A.S.; Buyya, R. Big Data Computing and Clouds: Trends and Future Directions. *J. Parallel Distrib. Comput.* **2015**, *79–80*, 3–15. [CrossRef]
- Kurniawati, R. Integrating Cloud Computing across Existing Computer and Information Science Courses. *J. Comput. Sci. Coll.* **2020**, *35*, 259–260.
- Deb, D.; Fuad, M. Integrating Big Data and Cloud Computing Topics into the Computing Curricula: A Modular Approach. *J. Parallel Distrib. Comput.* **2021**, *157*, 303–315. [CrossRef]
- Wang, L.; Ranjan, R.; Chen, J.; Benatallah, B. *Cloud Computing: Methodology, Systems, and Applications*; CRC Press: Boca Raton, FL, USA, 2017; ISBN 9781439856420.
- Nicolás-Agustín, Á.; Jiménez-Jiménez, D.; Maeso-Fernandez, F. The Role of Human Resource Practices in the Implementation of Digital Transformation. *Int. J. Manpow.* **2022**, *43*, 395–410. [CrossRef]
- Costello, K. Gartner Forecasts Worldwide Public Cloud End-User Spending to Grow 18% in 2021. Available online: <https://www.gartner.com/en/newsroom/press-releases/2020-11-17-gartner-forecasts-worldwide-public-cloud-end-user-spending-to-grow-18-percent-in-2021> (accessed on 15 February 2022).
- Vaishya, R.; Haleem, A.; Vaish, A.; Javaid, M. Emerging Technologies to Combat the COVID-19 Pandemic. *J. Clin. Exp. Hepatol.* **2020**, *10*, 409–411. [CrossRef] [PubMed]
- Kratzke, N. A Brief History of Cloud Application Architectures. *Appl. Sci.* **2018**, *8*, 1368. [CrossRef]
- Gurcan, F.; Cagiltay, N.E. Big Data Software Engineering: Analysis of Knowledge Domains and Skill Sets Using LDA-Based Topic Modeling. *IEEE Access* **2019**, *7*, 82541–82552. [CrossRef]
- Gurcan, F.; Kose, C. Analysis of Software Engineering Industry Needs and Trends: Implications for Education. *Int. J. Eng. Educ.* **2017**, *33*, 1361–1368.
- De Mauro, A.; Greco, M.; Grimaldi, M.; Ritala, P. Human Resources for Big Data Professions: A Systematic Classification of Job Roles and Required Skill Sets. *Inf. Process. Manag.* **2018**, *54*, 807–817. [CrossRef]
- Gurcan, F.; Sevik, S. Business Intelligence and Analytics: An Understanding of the Industry Needs for Domain-Specific Competencies. In Proceedings of the 1st International Informatics and Software Engineering Conference: Innovative Technologies for Digital Transformation, IISEC 2019-Proceedings, Ankara, Turkey, 6–7 November 2019.
- Gardiner, A.; Aasheim, C.; Rutner, P.; Williams, S. Skill Requirements in Big Data: A Content Analysis of Job Advertisements. *J. Comput. Inf. Syst.* **2018**, *58*, 374–384. [CrossRef]
- Yang, Q.; Zhang, X.; Du, X.; Bielefeld, A.; Liu, Y.Q. Current Market Demand for Core Competencies of Librarianship-A Text Mining Study of American Library Association's Advertisements from 2009 through 2014. *Appl. Sci.* **2016**, *6*, 48. [CrossRef]

21. Debortoli, S.; Müller, O.; Vom Brocke, J. Comparing Business Intelligence and Big Data Skills: A Text Mining Study Using Job Advertisements. *Bus. Inf. Syst. Eng.* **2014**, *6*, 289–300. [CrossRef]
22. Gurcan, F. Extraction of Core Competencies for Big Data: Implications for Competency-Based Engineering Education. *Int. J. Eng. Educ.* **2019**, *35*, 1110–1115.
23. Debaio, D.; Yinxia, M.; Min, Z. Analysis of Big Data Job Requirements Based on K-Means Text Clustering in China. *PLoS ONE* **2021**, *16*, e0255419. [CrossRef]
24. Adams, J.; Hainey, B.; White, L.; Foster, D.; Hall, N.; Hills, M.; Hooshangi, S.; Kuber, K.; Nazir, S.; Sakr, M.; et al. Cloud Computing Curriculum: Developing Exemplar Modules for General Course Inclusion. In Proceedings of the Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE, Online, 15 June 2020; pp. 510–511.
25. Foster, D.; White, L.; Erdil, D.C.; Adams, J.; Argüelles, A.; Hainey, B.; Hyman, H.; Lewis, G.; Nazir, S.; Nguyen, V.; et al. Toward a Cloud Computing Learning Community. In Proceedings of the Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE, Online, 18 December 2019; pp. 143–155.
26. Foster, D.; White, L.; Adams, J.; Cenker Erdil, D.; Hyman, H.; Kurkovsky, S.; Sakr, M.; Stott, L. Cloud Computing: Developing Contemporary Computer Science Curriculum for a Cloud-First Future. In Proceedings of the Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE, Online, 2 July 2018; pp. 130–147.
27. Indeed Employment. Indeed.Com. Available online: <https://www.indeed.com> (accessed on 15 February 2022).
28. Indeed Cloud Jobs, Employment. Indeed.Com. Available online: <https://www.indeed.com/jobs?q=title%3Acloud&vjk> (accessed on 15 February 2022).
29. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
30. Blei, D.M. Probabilistic Topic Models. *Commun. ACM* **2012**, *55*, 77–84. [CrossRef]
31. Gurcan, F.; Cagiltay, N.E.; Cagiltay, K. Mapping Human–Computer Interaction Research Themes and Trends from Its Existence to Today: A Topic Modeling-Based Review of Past 60 Years. *Int. J. Hum. Comput. Interact.* **2021**, *37*, 267–280. [CrossRef]
32. Gurcan, F.; Dalveren, G.G.M.; Cagiltay, N.E.; Soylu, A. Detecting Latent Topics and Trends in Software Engineering Research Since 1980 Using Probabilistic Topic Modeling. *IEEE Access* **2022**, *10*, 74638–74654. [CrossRef]
33. Konrad, M. Text Mining and Topic Modeling Toolkit. Available online: <https://pypi.org/project/tmtoolkit/> (accessed on 21 January 2022).
34. Gurcan, F.; Ozyurt, O.; Cagiltay, N.E. Investigation of Emerging Trends in the E-Learning Field Using Latent Dirichlet Allocation. *Int. Rev. Res. Open Distrib. Learn.* **2021**, *22*, 1–18. [CrossRef]
35. Gurcan, F.; Cagiltay, N.E. Exploratory Analysis of Topic Interests and Their Evolution in Bioinformatics Research Using Semantic Text Mining and Probabilistic Topic Modeling. *IEEE Access* **2022**, *10*, 31480–31493. [CrossRef]
36. Mimno, D.; Wallach, H.M.; Talley, E.; Leenders, M.; McCallum, A. Optimizing Semantic Coherence in Topic Models. In Proceedings of the EMNLP 2011-Conference on Empirical Methods in Natural Language Processing, Online, 27 July 2011.
37. Breivold, H.P.; Crnkovic, I. Cloud Computing Education Strategies. In Proceedings of the 2014 IEEE 27th Conference on Software Engineering Education and Training, CSEE and T 2014-Proceedings, Klagenfurt, KAE, Austria, 23–25 April 2014; pp. 29–38.
38. Bermbach, D.; Chandra, A.; Krintz, C.; Gokhale, A.; Slominski, A.; Thamsen, L.; Cavalcante, E.; Guo, T.; Brandic, I.; Wolski, R. On the Future of Cloud Engineering. In Proceedings of the Proceedings-2021 IEEE International Conference on Cloud Engineering, IC2E 2021, Online, 4–8 October 2021; pp. 264–275.
39. Hu, F.; Qiu, M.; Li, J.; Grant, T.; Tylor, D.; McCaleb, S.; Butler, L.; Hamner, R. A Review on Cloud Computing: Design Challenges in Architecture and Security. *J. Comput. Inf. Technol.* **2011**, *19*, 25–55. [CrossRef]
40. Petrovic, J.; Pale, P. Achieving Scalability and Interactivity in a Communication Skills Course for Undergraduate Engineering Students. *IEEE Trans. Educ.* **2021**, *64*, 413–422. [CrossRef]
41. Jeganathan, L.; Khan, A.N.; Kannan Raju, J.; Narayanasamy, S. On a Frame Work of Curriculum for Engineering Education 4.0. In Proceedings of the 2018 World Engineering Education Forum-Global Engineering Deans Council, WEEF-GEDC 2018, Albuquerque, NM, USA, 12–16 November 2019.
42. Holik, I.; Sanda, I.D. The Possibilities of Improving Communication Skills in the Training of Engineering Students. *Int. J. Eng. Pedagog.* **2020**, *10*, 20–33. [CrossRef]