GÜSBD 2023; 12(2): 377 - 385 | **Gümüşhane Üniversitesi Sağlık Bilimleri Dergisi** | **Araştırma Makalesi**
GUJHS 2023; 12(2): 377 - 385 | Gümüşhane University Journal of Health Sciences | Original Article

Identification of Expertise Roles and Skill Sets Required for Careers in Health Information Management

Sağlık Bilgi Yönetiminde Kariyer İçin Gerekli Uzmanlık Rollerinin ve Beceri Setlerinin Tespit Edilmesi

Fatih GÜRCAN[1]

ABSTRACT

This study aims to investigate the competency areas and skill sets demanded on the job market for health information management (HIM), which plays a vital role in sustaining and enhancing the quality and efficacy of health services. In accordance with this objective, a semantic content analysis was performed on online HIM job postings using a quantitative method based on text mining and probabilistic topic modeling to identify the expertise roles and skill sets as semantic topics. Our findings revealed ten expertise roles and twenty-four skills that represent a broad spectrum of HIM professions' competency requirements. "Specialist" (17.57%), "Director" (17.05%), "Manager" (13.18%), "Coder" (12.40%), and "Technician" (11.11%) are the top five expertise roles for HIM. A competency taxonomy was developed for HIM professions based on the knowledge and skills revealed by 24 topics using topic modeling analysis. The HIM competencies were categorized as "Medical Knowledge" (39.92%), "Management Skills" (29.80%), "IT Skills" (16.09%), and "Soft Skills" (14.18%). Our findings may have significant implications for HIM candidates and professionals, healthcare industries, and academic institutions in their efforts to comprehend, evaluate, and develop the necessary competencies and skills for HIM careers.

Keywords: Health Information Management, Competencies and Skills, Human Resources Management, Machine Learning, Topic Modeling

ÖZ

Bu çalışma, sağlık hizmetlerinin kalitesinin ve etkinliğinin sürdürülmesinde ve arttırılmasında hayati bir rol oynayan sağlık bilgi yönetimi (SBY) için iş piyasasında talep edilen yetkinlik alanlarını ve beceri setlerini araştırmayı amaçlamaktadır. Bu amaç doğrultusunda, çevrimiçi SBY iş ilanları üzerinde, metin madenciliği ve olasılıksal konu modellemeye dayalı nicel bir yöntembilim kullanılarak, uzmanlık rollerini ve beceri setlerini anlamsal konular olarak belirlemek için anlamsal bir içerik analizi gerçekleştirilmiştir. Bulgularımız, SBY mesleklerinin yetkinlik gereksinimlerinin geniş bir yelpazesini temsil eden on uzmanlık rolü ve yirmi dört beceriyi ortaya koydu. "Uzman" (%17,57), "Direktör" (%17,05), "Menajer" (%13,18), "Kodlayıcı" (%12,40) ve "Teknisyen" (%11,11) SBY için ilk beş uzmanlık rolü olarak tespit edilmiştir. Konu modelleme analizi kullanılarak 24 konunun ortaya çıkardığı bilgi ve becerilere dayanarak SBY alanı için bir yetkinlik taksonomisi geliştirilmiştir. SBY yetkinlikleri "Sağlık Bilgisi" (%39,92), "Yönetim Becerileri" (%29,80), "Bilişim Becerileri" (%16,09) ve "Sosyal Beceriler" (%14,18) olarak kategorize edildi. Bulgularımız, SBY adayları ve uzmanları, sağlık kuruluşları ve akademik kurumların SBY yetkinliklerini ve becerilerini anlama, değerlendirme ve geliştirme çabaları için önemli faydalar sağlayabilir.

Anahtar Kelimeler: Sağlık Bilgi Yönetimi, Yetkinlikler ve Beceriler, İnsan Kaynakları Yönetimi, Makine Öğrenimi, Konu Modelleme

[1]Assit. Prof., Fatih GÜRCAN, Data Analytics, Karadeniz Technical University, Department of Management Information Systems, fgurcan@ktu.edu.tr, ORCID: 0000-0001-9915-6686

İletişim / Corresponding Author: Fatih GÜRCAN
e-posta/e-mail: fgurcan@ktu.edu.tr

Geliş Tarihi / Received: 02.04.2023
Kabul Tarihi/Accepted: 14.06.2023

INTRODUCTION

Health information management (HIM) includes the processes of acquiring, recording, analyzing, and maintaining digital and traditional health information vital to providing quality patient care. With the widespread digitalization of health records and procedures, traditional (document-based) records are being replaced by electronic health records (EHRs) or electronic medical records (EMRs).[1] Health information systems and technologies, and their derivative tools provide more contribution and efficiency to information management in the health industry. HIM is an information technology-oriented management structure applied to healthcare services.[2] HIM is a multidisciplinary field that combines business, health, medicine, science, management, and information technology. [3]

Health records and information are data about a patient's medical history, including symptoms, diagnoses, procedures, and outcomes.[1,3] A health record contains vital information such as a patient's history, laboratory results, assays, x-rays, clinical information, demographic information, epidemiological data, reference and coded health data, and treatments administered. [2]

Although EHR and EMR are often used interchangeably, there is a difference between them. EMR usually refers to a digitized chart of patients stored on a computer containing diagnoses, vaccination dates, and medications, while EHR is a digital record of a patient's health, including medical history, progress notes, and laboratory data.[4]

A patient's health information can be viewed individually for each patient to see how the patient's health has changed. In some cases, this individual patient information may also be considered as part of a larger dataset to understand how a population's health is changing and how medical interventions affect community health outcomes.[1,2,5]

HIM professionals plan information systems to support the information-intensive and information technology-based healthcare system, develop healthcare policy, and identify current and future information and organizational needs.[3] In addition, by effectively using information science and technologies, they manage the collection, storage, analysis, use, maintenance, and transmission of medical records to meet the legal, professional, ethical, and administrative record-keeping requirements of healthcare delivery and ensure an uninterrupted supply of these vital services. enable them to continue. HIM professionals have a direct impact on the quality of services provided at every stage of the healthcare delivery cycle.[6]

Health information management professionals organize, audit, and maintain patient records and corporate business information in large healthcare systems, physicians' offices, and many other healthcare organizations. They also enter patient information into computer databases to comply with federal legislation mandating the use of patients' EHRs or EMRs.[4]

As the use of health services and applications based on information technologies has become widespread, the need for qualified labor in the field of HIM has increased day by day, and HIM has become one of the most important employment sources for today's health industries.[2,3]

In this regard, new career opportunities are opening up for HIM professionals every day. In the life cycle of services and applications based on health information systems, HIM professionals may assume different responsibilities under different roles, depending on the scope and type of work given.[2] Analyzing and defining the roles (job titles and expertise roles) and responsibilities (competencies and skills) of HIM professionals can yield important implications for a better understanding of all available HIM professions and the core competencies needed. From this perspective, there is a clear research gap in identifying the

roles and related skills needed for HIM professions and defining the dimensions of academic supply to meet this need.[1]

This study aims to identify and categorize the competency requirements for HIM professionals from a practitioner's point of view, thereby filling this previously expressed knowledge gap. To achieve this, a content analysis was performed on HIM-related online job postings using the Latent Dirichlet Allocation (LDA),[7,8] a generative topic modeling approach. As a result of this analysis, ten expertise roles for HIM and 24 topics (skills) that define the competencies of these roles were discovered. The findings of this study may offer important insights for HIM communities with different profiles.

MATERIAL AND METHOD

The methodology of this research includes the stages of discovering the competencies needed in the HIM job market based on topic modeling processes. Using this methodology, a semantic content analysis based on probabilistic topic modeling was performed on the experimental dataset of online job postings related to HIM. The methodology of this study consists of three stages: data collection, data preprocessing, and topic modeling analysis.

Data Collection and Preprocessing

In order to create an optimal dataset for this study, we initially tried to identify an appropriate data source (an online job posting site) from which we could get up-to-date job postings related to HIM. Among these possible sources, Indeed.com, a popular employment site with extensive search, filtering, and retrieval options, was chosen as the data source for this study.[8] Indeed.com is the world's #1 job site, with over 250 million unique visitors each month.[9] In this context, we considered all job postings with the terms "health", "information", and "management" in the job title as HIM professions to include the most appropriate job postings in the dataset. Examples of job postings that make up our empirical dataset can be viewed on the Indeed website using our search query.[9] In this way, 774 HIM-oriented job postings published in the two-month period between January 20, 2023, and March 20, 2023 were extracted using an API offered by Indeed.com, and an experimental data set was created. Each job posting that makes up our dataset consists of two subsections containing the job title and job description.

After the experimental data set was created, data preprocessing steps, which are one of the most necessary tasks of text mining, were applied sequentially to the textual data.[8] Data preprocessing covers the processes required for structuring unstructured textual data, especially web contents.[10,11] Initially, the texts of the jobs were split into words (known as tokens), known as tokenization, to obtain meaningful attributes, and all texts were converted to lowercase. Punctuation, web links, html tags, and meaningless characters were then deleted. The English stop words (a, an, the, who, is, are, etc.) were removed from the texts. Finally, the lemmatization process was applied to the texts in order to reduce the words from their derived form to their nominative form.[12] Finally, the document-term matrix (DTM), which provides the numerical representation of the texts, was created with the remaining words.[13,14]

Topic Modeling Analysis

In the first stage of the empirical analysis, we developed a taxonomy for HIM expertise areas by categorizing job titles according to their key expertise roles. The job title of each HIM profession is an important indicator of the expertise role defined for that job, i.e., in which role the HIM expert will work. From this point of view, each title has been assigned to a role related to HIM expertise areas, taking into account the words in the job titles. For example, job titles such as "Director of Health Information Management" or "Health Information Management Director", which contain derivatives of the word "director", are assigned to the expertise area of the director

role.[8,15] Other words in HIM job titles were identified in this way, and all job titles were assigned to their respective expertise roles. In this way, ten expertise roles have been identified for HIM professions.

After determining the expertise roles, topic modeling analysis was performed on the experimental corpus of HIM job postings using the Latent Dirichlet allocation (LDA) algorithm.[7,16] In semantic text analysis, LDA topic modeling is a probabilistic approach used to model the semantic structure of a textual corpus.[17,18] Text documents contain hidden semantic patterns called "topics" and each of these topics represents a specific skill in job postings. A text document can contain multiple topics in varying proportions. This highlights the fact that a job posting can include many different skills at different rates. LDA is an unsupervised machine learning model that can discover hidden topics (skills in this study) from an unstructured corpus of text without any training or annotation.[19] Therefore, the LDA algorithm was used for the topic modeling analysis of the HIM professions.

To implement the LDA model in our HIM corpus, we used Gensim, a pure Python library for topic modeling.[20] In order to select the optimal number of topics (K), we applied the LDA model with different K values ranging from 15 to 50.[21] We calculated a coherence score for each adapted model for each K value changing in this range.[22] We evaluated the calculated coherence scores for each K in this range and obtained the maximum score for the number of K = 24 topics.[20] Next, we identified the top 15 descriptive keywords with the highest frequency for each of the 24 topics that best represented the HIM skills at the optimal semantic level. Then, a topic name was assigned to each topic, taking into account the descriptive keywords of the topics. In addition, the distribution of the discovered topics per job posting, the distribution of words per topic, and the distribution of the topics in the entire corpus were calculated.[7,20,23] As a result, these 24 topics discovered by LDA-based topic modeling analysis were considered a representation of HIM skills in all subsequent analyses.

RESULTS AND DISCUSSION

This section presents the results obtained from the analysis of HIM job postings. In this study, a semantic content analysis was conducted on online job postings using LDA-based topic modeling to explore the expertise roles, knowledge domains, and skill sets required for HIM. Through the analysis and interpretation of the skills discovered as topics with the LDA, the roles and responsibilities of HIM professionals were determined, as well as the interdisciplinary scope of HIM professions.

In Table 1, the top 20 job titles from all jobs are given in order of percentile volume. A total of 298 different job titles have been identified for 774 job postings in our dataset, and the first 20 titles constitute 55.84% of the total titles. The first three titles were determined as "Health Information Management Clerk" (4.91%), "Health Information Management Specialist" (4.39%), and "Director Health Information Management" (4.39%).

We then analyzed the most frequently used terms related to roles in job titles. Taking into account the frequency of these terms in job titles in our dataset, a total of ten expertise roles were determined for HIM professions from job postings.

These expertise roles and percentages are given in Table 2. As seen in Table 2, HIM professions contain a total of 10 expertise roles. The first five of these roles were identified as "Specialist" (17.57%), "Director" (17.05%), "Manager" (13.18%), "Coder" (12.40%), and "Technician" (11.11%). The ratio of the first five roles to the total jobs was calculated as 71.32%. In other words, close to three quarters of HIM professions referred to the top five expertise roles. The total ratio of "Instructor", "Supervisor", "Clerk", "Coordinator", and

"Analyst" roles for HIM professions is 28.68%. The expertise roles given in Table 2 make it clear that HIM professions require an interdisciplinary background and a wide range of competencies.

Table 1. Top 20 HIM Job Titles and Their Percentages

Job Title	Rate (%)
Health Information Management Clerk	4.91
Health Information Management Specialist	4.39
Director Health Information Management	4.39
Health Information Management Coordinator	4.13
Health Information Management Director	4.13
Health Information Management Technician	3.88
Director of Health Information Management	3.88
Manager Health Information Management	3.88
Specialist Health Information Management	3.36
Health Information Management Manager	2.84
Coding Specialist Health Information Management	2.58
Health Information Management	2.33
Health Information Management Analyst	2.33
Supervisor Health Information Management	2.07
Health Information Management Supervisor	1.55
Clinical Documentation Specialist Health Information Management	1.29
Coder II Health Information Management	1.29
Manager of Health Information Management	1.03
Health Information Management Assistant	0.78
Health Information Management Lead	0.78

Table 2. Expertise Roles for HIM Professions

Expertise Role	Rate (%)
Specialist	17.57
Director	17.05
Manager	13.18
Coder	12.40
Technician	11.11
Instructor	6.72
Supervisor	6.46
Clerk	6.20
Coordinator	4.91
Analyst	4.39

At this stage of our results, the findings from the topic modeling analysis based on LDA are presented. As a result of this topic modeling analysis, a total of 24 topics (skills) were discovered for HIM professions. These topics are presented in Table 3 in order of percentage. In Table 3, the first word for any topic is the most common word, and the last word is the word with the least frequency in that topic. These 24 topics, discovered through topic modeling analysis, can be considered the knowledge and skills needed in this field. In this context, the terms "topic" and "skill" are used interchangeably throughout the text. These topics reveal that HIM professionals must possess a wide range of knowledge and skill sets. Table 3 shows that the three most prominent skills are "Medical Records Management" (7.51%), "Communication Skills" (6.87%), and "Certification" (6.39%). On the other hand, the skills with the lowest percentages are "Project Management" (1.09%), "ROI (Return on Investment) Management" (1.78%), and "Employee Engagement" (1.80%), respectively.

Table 3. Topics (Skills) Discovered by LDA

Topic Name	Descriptive Keywords	%
Medical Records Management	medical patient record hospital management responsible physician care manage completion function standard regulatory quality prefer	7.51
Communication Skills	skill ability knowledge require demonstrate office computer write maintain communication customer communicate talk excellent terminology	6.87
Certification	experience certification rhia rhit degree require associate coding bachelor ahima level least preferred set field	6.39
Work Experience	work experience require need provide requirement practice knowledge assist schedule professional meet time perform activity	6.16
Information Services	information service system department procedure software technician register ehr emr regulation policy data provide privacy	6.16
Educational Requirements	education diploma job bachelor high school qualification shift position degree minimum equivalent primary full-time day	6.13
Performance Management	performance management operation ensure director leadership department process manage plan revenue improvement implement function budget	5.85
Compliance Management	compliance staff management program service coordinate education development review training employee serve develop organization facility	5.43
Medical Record Processing	record maintain medical report review file process necessary processing daily chart organize time accord duty	5.17
Hospital Management	director manager hospital management service team business full salary organization opportunity lead career officer responsible	5.16
Healthcare Quality	health care service quality satisfaction responsibility provide healthcare summary improve electronic support excellence mission recognize	4.71
Electronic Recording	record document electronic storage perform accuracy software scan timely process deficiency quality analysis assist completion	4.07
Medical Coding	code coder coding hospital outpatient inpatient guideline cc cpt procedure diagnosis terminology assign query certified	3.84
Medical Data Analytics	system data analytics report visual administrative experience procedure accordance analysis time unit healthcare information develop	3.25
Problem Solving	problem solve support work duty issue customer resolve technical perform change time maintain complete specialist	3.02
Documentation Management	health documentation management system clinical ensure oversee care document compliance experience analysis timely improvement hospital	2.99
Clinical Documentation	documentation clinical specialist physician care review reimbursement hospital opportunity appropriate quality case nurse specialty drg	2.77
Technological Know-How	technology service information digital technical knowledge opportunity equal job detail qualification documentation schedule excellent description	2.62
Patient Document Requests	request patient document emr record chart scan image complete department copy fax log verify birth	2.51
Team Working	team professional join working healthcare help teamwork want knowledge lead seek get application grow value	2.49
Inpatient Rehab Services	rehabilitation inpatient hospital service facility care acute term experience offer injury uh disability help treatment	2.25
Employee Engagement	engage employee involve volunteer motivation commitment service provider people management individual call employer network deliver	1.80
ROI Management	manager roi retrieve investment registry management training performance applicable employee participant assignment process work field	1.78
Project Management	project data management team research plan collection opportunity technical presentation report ehr development support evaluation	1.09

In the next step of our analysis, the topics are categorized in order to provide a more understandable presentation of the knowledge and skills required for HIM. For this purpose, a systematic categorization was carried out by associating knowledge and skills with their relevant backgrounds, competency areas, and workflows. HIM knowledge and skills, represented by 24 topics, were classified under four main categories, and a competency taxonomy was developed for HIM. This competency taxonomy is presented in Table 4 with percentages. The first category, "Medical Knowledge" (39.92%), includes nine skill items consisting of "Certification", "Work Experience", "Educational Requirements", "Medical Record Processing", "Healthcare Quality", "Medical Coding", "Clinical Documentation", "Patient Document Requests", and "Inpatient Rehab Services". The second category, "Management Skills" (29.80%), contains seven skill items consisting of "Medical Records Management", "Performance Management", "Compliance Management", "Hospital Management", "Documentation Management", "ROI Management", and "Project Management". The third category is "IT Skills" (16.09%) and includes four items of skills "Information Services", "Electronic Recording", "Medical Data Analytics", and "Technological Know-How". The fourth category is "Soft Skill" (14.18%) and this category includes four items of skills "Communication Skills", "Problem Solving", "Team Working", and "Employee Engagement".

Table 4. Taxonomy of the HIM Competencies and Skills

Category	Topics (Skills)	Rate (%)	Total (%)
Medical Knowledge	Certification	6.39	39.92
	Work Experience	6.16	
	Educational Requirements	6.13	
	Medical Record Processing	5.17	
	Healthcare Quality	4.71	
	Medical Coding	3.84	
	Clinical Documentation	2.77	
	Patient Document Requests	2.51	
	Inpatient Rehab Services	2.25	
Management Skills	Medical Records Management	7.51	29.80
	Performance Management	5.85	
	Compliance Management	5.43	
	Hospital Management	5.16	
	Documentation Management	2.99	
	ROI Management	1.78	
	Project Management	1.09	
IT Skills	Information Services	6.16	16.09
	Electronic Recording	4.07	
	Medical Data Analytics	3.25	
	Technological Know-How	2.62	
Soft Skills	Communication Skills	6.87	14.18
	Problem Solving	3.02	
	Team Working	2.49	
	Employee Engagement	1.80	

Our findings highlighted the importance of competent workforce resources for the HIM field, involving dynamic work environments where tasks, paradigms, technologies, skills and experiences are constantly changing and evolving.[2] Our analysis revealed the competencies demanded for HIM professionals and their relationships to expertise roles. Our findings also revealed the competency areas and skill sets needed in the HIM field in the form of 24 different topics generated by the LDA. We now discuss these findings in detail.[1,2]

Our findings highlighted a wide spectrum of competencies needed in the ever-growing HIM industry, from communication skills to management skills, technical knowledge-domains to information technology (IT) skills, medical coding to documentation systems. According to our findings, social skills are also highly demanded for HIM professionals. Our findings on soft skills include the topics of "Communication Skills", "Problem Solving", "Team Working", and "Employee Engagement" (see Table 4).[2] Among the discovered topics, "Medical Records Management" (7.51%) is the top skill, demonstrating how important medical records management and documentation are for HIM expertise. The importance of medical record and documentation management for HIM has been emphasized in previous studies. [3,4]

Another important finding of our study is that a qualified educational background is necessary for HIM professionals.[24] More specifically, the topics "Certification" (6.39%), "Work Experience" (6.16%), and "Educational Requirements", (6.13%) emphasize the necessity of experience and educational qualifications for HIM expertise. The total ratio of these three topics is 18.68%, indicating the dominant role of educational and experiential competencies among all topics.[24]

Topics such as "Medical Record Processing", "Electronic Recording", and "Medical Coding" that we discovered through our analysis revealed that traditional medical records are replaced by electronic health records (EHR) or electronic medical records (EMR). Based on these findings, we can say that the HIM field is evolving more and more towards a management model focused on electronic data and digital documentation.[3–5,25]

CONCLUSION AND RECOMMENDATIONS

This study is based on a semantic content analysis of HIM-related job postings using text mining and probabilistic topic modeling techniques to identify the competencies, knowledge domains, and skill sets required by the HIM industry. This study found that (1) the expertise roles and skills explored are also emerging trends and demands in HIM work environments and qualifications that HIM professionals must possess; (2) HIM expertise roles require a broad set of competencies with an interdisciplinary background; (3) HIM professions contain many wide-ranging competencies (24 skills) consisting of medical, management, IT, and social skills; and (4) Careers in HIM include ten expertise roles and 298 job titles. Since there is no experimental study in this context in the existing literature, it is anticipated that this study will contribute significantly to the HIM communities. This study's findings can provide valuable insight into the fundamental structure and competency requirements of HIM professions. Our findings can assist HIM professionals in measuring and updating their competencies, trainers in preparing HIM candidates to meet emerging market demands, and students in planning their careers. In addition, this study's methodology can be used in future studies to reveal the needs and trends of various industries.

REFERENCES

1. Fenton, S.H, Low, S, Abrams, K.J. and Butler-Henderson, K. (2017). "Health Information Management: Changing with Time". In Yearbook of Medical Informatics 26 (1), 72–77. https://doi.org/10.15265/IY-2017-021

2. Beesley, K, McLeod, A, Hewitt, B. and Moczygemba, J. (2021). "Health Information Management Reimagined: Assessing Current Professional Skills and Industry Demand". Perspectives in Health Information Management, 18 (Winter).

3. Berg, M. (2004). "Health Information Management: Integrating Information Technology in Health Care Work". Psychology Press.

4. Ehrenstein, V, Kharrazi, H, Lehmann, H. and Taylor, C.O. (2019). "Obtaining Data from Electronic Health Records". Tools and Technologies for Registry Interoperability, Registries for Evaluating Patient Outcomes: A User's Guide, 1–92.

5. Gurcan, F, Boztas, G.D, Dalveren, G.G.M. and Derawi, M. (2023). "Digital Transformation Strategies, Practices, and Trends: A Large-Scale Retrospective Study Based on Machine Learning". Sustainability, 15 (9), 7496. https://doi.org/10.3390/su15097496

6. Wager, K.A, Lee, F.W. and Glaser, J.P. (2021). "Health Care Information Systems: A Practical Approach for Health Care Management". John Wiley and Sons.

7. Blei, D.M. (2012). "Probabilistic Topic Models". Communications of the ACM, 55 (4), 77–84.

8. Ozyurt, O, Gurcan, F, Dalveren, G.G.M. and Derawi, M. (2022). "Career in Cloud Computing: Exploratory Analysis of In-Demand Competency Areas and Skill Sets". Applied Sciences, 12 (19), 9787.

9. Indeed. (2023). Job Search | Indeed. https://www.indeed.com/

10. Gurcan, F, Dalveren, G.G.M, Cagiltay, N. E. and Soylu, A. (2022). "Detecting Latent Topics and Trends in Software Engineering Research Since 1980 Using Probabilistic Topic Modeling". IEEE Access, 10, 74638–74654. https://doi.org/10.1109/ACCESS.2022.3190632

11. Gurcan, F. and Cagiltay, N.E. (2022). "Exploratory Analysis of Topic Interests and Their Evolution in Bioinformatics Research Using Semantic Text Mining and Probabilistic Topic Modeling". IEEE Access, 10, 31480–31493. https://doi.org/10.1109/ACCESS.2022.3160795

12. Srivastava, A.N. and Sahami, M. (2009). "Text Mining: Classification, Clustering, and Applications". CRC Press.

13. Karl, A, Wisnowski, J. and Rushing, W.H. (2015). "A Practical Guide to Text Mining with Topic Extraction". Wiley Interdisciplinary Reviews: Computational Statistics, 7 (5), 326–340.

14. Gurcan, F, Dalveren, G.G.M, Cagiltay, N.E, Roman, D. and Soylu, A. (2022). "Evolution of Software Testing Strategies and Trends: Semantic Content Analysis of Software Research Corpus of the Last 40 Years". IEEE Access, 10, 106093–106109. https://doi.org/10.1109/ACCESS.2022.3211949

15. Gurcan, F. (2019). "Extraction of Core Competencies for Big Data: Implications for Competency-Based Engineering Education". International Journal of Engineering Education, 35 (4), 1110–1115.

16. Blei, D.M, Ng, A.Y. and Jordan, M.I. (2003). "Latent Dirichlet Allocation". Journal of Machine Learning Research, 3 (4/5), 993–1022.

17. Gurcan, F, Cagiltay, N.E. and Cagiltay, K. (2021). "Mapping Human–Computer Interaction Research Themes and Trends from Its Existence to Today: A Topic Modeling-Based Review of past 60 Years". International Journal of Human-Computer Interaction, 37 (3), 267–280. https://doi.org/10.1080/10447318.2020.1819668

18. Blei, D.M. and Lafferty, J.D. (2007). "Correction: A Correlated Topic Model of Science". The Annals of Applied Statistics, 1 (2), 634–634. https://doi.org/10.1214/07-aoas136

19. Vayansky, I. and Kumar, S.A.P. (2020). "A Review of Topic Modeling Methods". Information Systems, 94. https://doi.org/10.1016/j.is.2020.101582

20. Řehůřek, R. and Sojka, P. (2011). "Gensim—Statistical Semantics in Python". In Lecture Notes in Computer Science.

21. Röder, M, Both, A. and Hinneburg, A. (2015). "Exploring The Space of Topic Coherence Measures". WSDM 2015 - Proceedings of the 8th ACM International Conference on Web Search and Data Mining. https://doi.org/10.1145/2684822.2685324

22. Mimno, D, Wallach, H. M, Talley, E, Leenders, M. and McCallum, A. (2011). "Optimizing Semantic Coherence in Topic Models". EMNLP 2011 - Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference.

23. Gurcan, F. and Ozyurt, O. (2020). "Emerging Trends and Knowledge Domains in E-Learning Researches: Topic Modeling Analysis with the Article Published between 2008-2018". Journal of Computer and Education Research, 8 (16), 738–756. https://doi.org/10.18009/jcer.769349

24. Bates, M, Black, C, Blair, F, Davis, L, Ingram, S, Lane, D. Q, McElderry, A, Peagler, B, Pickett, J, Plettenberg, C. and Hart-Hester, S. (2014). "Perceptions of Health Information Management Educational and Practice Experiences". Perspectives in Health Information Management / AHIMA, American Health Information Management Association, 11.

25. Gurcan, F. (2023). "What Issues Are Data Scientists Talking About? Identification of Current Data Science Issues Using Semantic Content Analysis of Q&A Communities". PeerJ Computer Science, 9, e1361. https://doi.org/10.7717/peerj-cs.1361