Gümüşhane University Journal of Science

GUFBD / *GUJS* (2025) 15(1): 197-209 doi: 10.17714/gumusfenbil.1600286 Research Article

Next-generation software development competencies: Identification of technical and non-technical skills needed by modern industry

Yeni nesil yazılım geliştirme yetkinlikleri: Modern endüstrinin ihtiyaç duyduğu teknik ve teknik olmayan becerilerin belirlenmesi

Fatih GÜRCAN*[1], **Cemal KÖSE**[2]

[1]*Karadeniz Technical University, Department of Management Information Systems, 6080, Trabzon, Turkey*

[2]*Karadeniz Technical University, Department of Computer Engineering, 61080, Trabzon, Turkey*

• Received: 12.12.2024 • Accepted: 14.02.2025

Abstract

The software development industry is undergoing unprecedented growth and transformation, prompting a reevaluation of the skills and competencies necessary for success in this dynamic landscape. This study investigates the rapidly evolving skill requirements within the industry, driven by technological advancements. To achieve this, a Latent Dirichlet Allocation (LDA) framework is employed, enabling the identification of key topics from a dataset derived from online job postings. The analysis revealed 52 core topics pertinent to software development competencies. The findings reveal a significant emphasis on both technical domain knowledge and programming skills, with particular attention to modern programming languages such as Java, Python, and JavaScript. Moreover, non-technical skills, including communication, teamwork, and critical thinking, are underscored as vital competencies in today's collaborative software development environments. These insights emphasize the necessity for software developers to cultivate a diverse skill set to adapt to current and future industry demands. This work serves as a crucial reference for understanding the present and future skill requirements in the software development field, providing valuable guidance for developers, employers, and educational institutions.

Keywords: Competency analysis, Skill requirements, Software development skills, Topic modeling

Öz

Yazılım geliştirme sektörü önemli bir büyüme ve dönüşümden geçiyor ve bu dinamik ortamda başarı için gerekli becerilerin ve yeterliliklerin yeniden değerlendirilmesi gerekiyor. Bu çalışma, teknolojik gelişmelerin yönlendirdiği sektördeki hızla değişen beceri gereksinimlerini araştırıyor. Bunu başarmak için, çevrimiçi iş ilanlarından türetilen bir veri setinden temel konuların belirlenmesini sağlayan bir Gizli Dirichlet Tahsisi (Latent Dirichlet Allocation) çerçevesi kullanılıyor. Gerçekleştirilen analiz, yazılım geliştirme yetkinliklerine özgü 52 temel beceriyi ortaya çıkarmıştır. Bulgular, Java, Python ve JavaScript gibi modern programlama dillerine özel dikkat gösterilerek hem teknik alan bilgisine hem de programlama becerilerinin önemine vurgu yapmıştır. Dahası, iletişim, ekip çalışması ve eleştirel düşünme gibi teknik olmayan beceriler, günümüzün işbirlikçi yazılım geliştirme ortamlarında çok önemli yetkinlikler ortaya çıkmıştır. Bu içgörüler, yazılım geliştiricilerinin mevcut ve gelecekteki endüstri taleplerine uyum sağlamak için çeşitli bir beceri seti geliştirmeleri gerekliliğini vurgulamaktadır. Bu çalışma yazılım geliştirme alanındaki mevcut ve gelecekteki beceri gereksinimlerini anlamak için önemli bir referans görevi görerek geliştiriciler, işverenler ve eğitim kurumları için değerli rehberlik sağlamaktadır.

Anahtar kelimeler: *Beceri gereksinimleri, Konu modelleme, Yazılım geliştirme becerileri, Yetkinlik analizi*

*Fatih GÜRCAN; fgurcan@ktu.edu.tr

ISSN: 2146-538X e-ISSN: 2146-538X https://dergipark.org.tr/tr/pub/gumusfenbil

1. Introduction

The software industry has undergone significant transformation in recent years due to rapidly evolving technologies and increasing competition. Modern software development processes now require a multidisciplinary skill set that extends beyond technical knowledge and expertise. In this dynamic environment, software developers must not only enhance their technical proficiency but also develop competencies in areas such as analytical thinking, problem-solving, teamwork, and communication. As technology advances, the demand for software development professionals is rapidly changing, with evolving needs for programming languages and software development tools (Aken et al., 2010; Debortoli et al., 2014; Gurcan, 2023a). In this context, understanding the latest trends in the software industry and identifying the skills and knowledge required to meet these demands is crucial for both employers shaping the sector and developers advancing their careers (Gurcan & Cagiltay, 2019).

Existing studies indicate that skill requirements in the software development field are evolving rapidly, and the competencies needed by software professionals continue to expand (Aken et al., 2010; Thomas et al., 2014; Terblanche & Wongthongtham, 2015; Chen et al., 2016; Vayansky & Kumar, 2020). A series of studies have highlighted that with the increasing pace of technological innovations, developers need more than just technical knowledge; they also require soft skills such as analytical thinking, problem-solving, and teamwork (Aken et al., 2010; Moreno et al., 2012; Debortoli et al., 2014; Gurcan & Cagiltay, 2019). Furthermore, as software projects become more complex, multiple studies have addressed the need for developers to be proficient in various programming languages and to quickly adapt to new tools (Moreno et al., 2012; Terblanche & Wongthongtham, 2015; Gurcan & Cagiltay, 2019). Probabilistic topic modeling techniques are frequently employed to analyze these dynamic industrial skill requirements and to identify the expertise demanded of software developers by examining large datasets from online job postings (Gurcan & Kose, 2017; De Mauro et al., 2018; Gurcan & Cagiltay, 2019). In these studies, methods such as Latent Dirichlet Allocation (LDA) help uncover hidden topics in text, revealing which skills are in greater demand and which techniques are gaining importance in the industry (Blei et al., 2003; Blei, 2012). Particularly, changes in next-generation programming languages, software development tools, and methodologies can be better understood through these analyses, aiding developers in planning their careers (Barua et al., 2014; De Mauro et al., 2018; Gurcan, 2023a).

This paper presents an analysis aimed at understanding the impact of technological transformation and innovation in the software development industry on the workforce, specifically focusing on the current demands for software development expertise. The primary goal of the study is to identify the expert domains, technical skills, competencies, and the necessary tools that form the foundation of the software development process in continually evolving environments. Accordingly, a comprehensive analysis was conducted on a large dataset of online job postings for software developers using the probabilistic topic modeling method based on LDA (Latent Dirichlet Allocation). Probabilistic topic modeling is a powerful tool for discovering hidden semantic structures in large datasets. In this study, the text data gathered from online job postings was analyzed using the LDA algorithm to uncover the most relevant knowledge and skills related to software development expertise. The 52 core topics identified were categorized into domains of expertise and technical knowledge requirements to highlight current trends in the software industry. Thus, the crucial knowledge and skills for software development processes were classified into four main competency areas: technical knowledge and skills (23%), software development and programming skills (52%), analytical and logical skills (12%), and non-technical skills (13%).

The findings demonstrate that the knowledge and skills required for software developers span a broad spectrum, with these trends becoming increasingly interdisciplinary. In particular, the importance of non-technical skills such as problem-solving abilities, teamwork, communication skills, and leadership, alongside technical proficiency, has been emphasized. In this context, software developers must not only possess coding skills but also be capable of collaborating effectively, working within teams, and making strategic decisions. The results presented in this paper indicate that the competencies required in software development processes have become multifaceted, and the development of these skills is critical for success in the industry. The findings provide a valuable guide for software professionals, helping them adapt to changes in the sector and shape their career development accordingly. This study contributes to identifying new trends in the software industry, enabling professionals to acquire the knowledge and skills needed for future demands. Additionally, by offering a taxonomy of both technical and non-technical skills in demand, it serves as a useful resource for educational institutions and employers.

2. Material and methods

The architecture of the study involves the stages of semantic analysis using an LDA-based topic modeling approach on a text-based qualitative dataset. In the first stage, an experimental dataset was created using data obtained from online job platforms. Subsequently, preprocessing steps were applied to the dataset to reduce dimensionality and enhance the accuracy of the analysis. In the next stage, the qualitative dataset was converted into a Document-Term Matrix (DTM) to enable numerical analyses. Following this step, a semantic analysis based on an LDA-based probabilistic topic modeling approach was performed on the DTM to uncover latent semantic structures. The topics derived from the experimental analysis were then discussed in terms of their semantic relevance, and the experimental analysis was reconfigured with different parameters until the desired level of semantic coherence was achieved. In the final stage, the results and interpretations derived from the analysis were presented and discussed in light of similar studies.

2.1. Data collection

In today's technological environment, online job platforms are widely used by both employers and employees (Debortoli et al., 2014; Gurcan & Cagiltay, 2019). These platforms serve as essential communication tools facilitating interactions between employers and job seekers (Gurcan & Cagiltay, 2019). As part of this study, several online job platforms were examined to determine a suitable data source. Two important criteria were considered in selecting the data source for the experimental analysis: first, the selected platform needed to be focused heavily on software development roles, and second, it needed to be a global platform featuring job postings for software development roles across different countries.

Considering these criteria, Stack Overflow's job platform (Stack Overflow Developer Jobs) was chosen as the data source (Stack Overflow, 2024). The job postings on Stack Overflow are specifically focused on software development roles (developer jobs) (Barua et al., 2014; Stack Overflow, 2024). Besides being a significant platform in the software development industry, Stack Overflow is also an important knowledge-sharing platform where developers interact through a question-and-answer format (Gurcan, 2023a). In this respect, Stack Overflow holds strategic importance for software developers (Barua et al., 2014; Stack Overflow, 2024). The dataset created for this study consists of 4,128 job postings focused on software development. Each record in the dataset typically includes information such as job title, date, location, job description, required knowledge, skills, qualifications, and competencies.

2.2. Data preprocessing

Text preprocessing is a critical step that directly impacts the performance of text-based data analyses (Gurcan, 2023b). The preprocessing procedures applied in this study included tokenization, text cleaning, removal of incomplete or irrelevant text entries, elimination of stop words, and frequency reduction, among other sequential processes (Gurcan, 2023b; Wu et al., 2024). Depending on the type of data being processed and the nature of the experimental analysis, preprocessing steps may vary. Given the high technical jargon content of the texts in this study's dataset, stemming was not applied to these texts (Chen et al., 2016).

Following these steps, terms with a frequency of fewer than five occurrences (i.e., words with negligible semantic contribution) were removed from the word space, further reducing the dimensionality of the dataset (Chen et al., 2016). This completed the preprocessing phase. The resulting word vectors for each text in the dataset were then compiled into a DTM (Burkhardt & Kramer, 2019). Once the preprocessing phase was completed, each text in the dataset was represented as a word vector. In the vector space model used to model the distribution of terms across documents, all numerical analyses were performed on these word vectors (Feng et al., 2022). To model the entire dataset, the term vectors for each document were combined into a single DTM.

2.3. Topic modeling analysis using LDA

Probabilistic topic modeling is a generative and probabilistic approach used to model the semantic structure of text documents within large collections and uncover latent semantic patterns. Text documents contain latent semantic structures known as topics (Blei et al., 2003; Blei, 2012). Each topic is characterized by a probability distribution over a fixed vocabulary of words. According to the probabilistic topic modeling approach, a text document can contain multiple topics with varying proportions (Barua et al., 2014; Gurcan & Cagiltay, 2019). The determination of these topic proportions forms the core of the probabilistic topic modeling approach.

In this study, we chose the Latent Dirichlet Allocation (LDA) method for topic modeling due to its proven effectiveness in discovering latent topics within large text datasets. LDA is widely used in text mining and provides a balance between interpretability and computational efficiency, making it suitable for our dataset (Blei et al., 2003; Blei, 2012). Although deep learning-based methods such as BERT (Bidirectional Encoder Representations from Transformers) and GPT (Generative Pretrained Transformers) have shown significant advancements in natural language processing, offering more contextual understanding, we focused on LDA for its established suitability in topic modeling. These alternative methods could be explored in future research for a more detailed contextual analysis and to provide a comparison of their effectiveness in this domain (Barua et al., 2014; Gurcan & Cagiltay, 2019).

LDA (Latent Dirichlet Allocation) is an unsupervised learning method that can be effectively applied to large text collections without requiring a training set. Since this method does not require prior learning, large textual datasets can be analyzed in a short time (Xu et al., 2022; Gurcan, 2023b). The LDA probabilistic topic modeling algorithm involves a series of matrix operations performed on the DTM to reveal semantic relationships between words (Egger & Yu, 2022). In this process, words with high co-occurrence rates are grouped, with each word group representing a “topic”. The process of extracting semantic topics from documents using the LDA approach is illustrated in Figure 1.

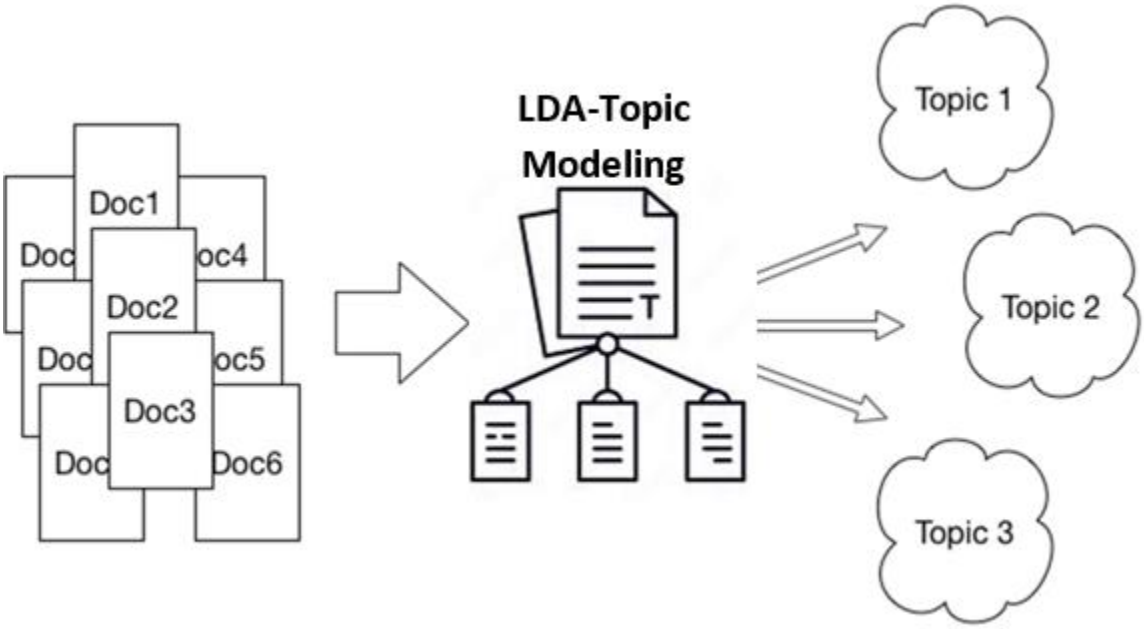

Figure 1. Extracting topics from documents using LDA

In LDA algorithms, Gibbs sampling is commonly used for parameter estimation. For the experimental analyses in this study, the MALLET tool (McCallum, 2002), which implements a scalable version of the LDA model based on Gibbs sampling, was used. MALLET, widely utilized in topic modeling applications, is a fast and scalable Java-based tool that employs Gibbs sampling for topic modeling (Blei et al., 2003). In the analysis performed using MALLET, the LDA algorithm was applied to the dataset with 1,000 iterations of Gibbs sampling, resulting in the identification of 52 meaningful topics. The number of topics for the LDA model was determined using coherence scores, which measure the semantic similarity of words within topics. Through an iterative evaluation process, we identified 52 topics as the optimal number, balancing coherence and interpretability.

3. Results and discussion

In this article, an experimental study was conducted using an LDA-based probabilistic topic modeling approach to identify expert knowledge and skills related to emerging trends in software development. The analysis was performed on a dataset derived from job postings for software developers. As a result of the topic modeling analysis on the experimental dataset, 52 topics were identified, revealing the most significant knowledge and skills associated with software development trends. These 52 topics, which represent trends in software development expertise, were organized into a taxonomy related to core disciplines and areas of software development expertise. Through this process, four main competency areas for expert knowledge and skills in software development trends were identified. Subsequently, the identified 52 topics were assigned to these areas based on their content, considering the keywords and percentage distributions within each topic. The percentage distribution of topics across the four main competency areas is as follows: technical knowledge and skills (23%), software development and programming skills (52%), analytical and logical skills (12%), and non-technical skills (13%). Additionally, the distribution of topics across these competency areas in terms of topic count is as follows: 13 topics for technical knowledge and skills, 22 topics for software development and programming skills, 10 topics for analytical and logical skills, and 7 topics for non-technical skills.

At first glance, the findings indicate that the knowledge and skills required in the software development industry encompass a broad range of disciplines. According to the results, software development expertise demands not only comprehensive software development skills but also technical domain knowledge, analytical and logical abilities, and individual (non-technical) skills such as communication. To provide a more detailed understanding of the knowledge and skills associated with emerging software development trends, and to further elaborate on the identified topics and competency areas, the findings are presented under four subheadings.

3.1. Technical knowledge and skills

Among the 52 identified topics, 13 were observed to fall under the “technical knowledge and skills” category, with a total frequency percentage of 23%. These 13 topics, assigned to the technical knowledge and skills area, are presented in Figure 2 along with their assigned topic names and percentage values. The findings related to technical knowledge and skills emphasize the theoretical knowledge areas that are essential for software expertise.

The most prominent topic is “Educational requirements” (4.90%), which indicates the strong emphasis on formal qualifications and foundational knowledge as key components in the competency framework. Following this, “Database knowledge” (2.70%) and “Deep domain knowledge” (2.60%) are highly emphasized, reflecting the importance of expertise in managing structured data and specialized knowledge in specific domains. “Software architectures” (2.10%) underscores the significance of designing robust and scalable systems, while “Software engineering knowledge” (1.80%) and “Software development processes” (1.70%) highlight the need for understanding software lifecycle management and best practices.

Topics such as “Big data processing” (1.60%), “Scaling systems” (1.20%), and “Cloud-based systems” (1.20%) represent emerging technical areas, focusing on the ability to handle large-scale data and scalable infrastructures. These reflect the growing reliance on cloud technologies and distributed systems in modern applications. Lesser-emphasized topics, including “Business intelligence” (0.90%), “Artificial intelligence” (0.80%), “Data security and privacy” (0.80%), and “System administration” (0.70%), indicate specialized areas of interest that, while not as dominant, are still essential for specific roles and applications in the field.

Overall, Figure 2 demonstrates a balanced distribution of foundational and advanced competencies, with a clear focus on formal education, domain-specific expertise, and modern technical proficiencies. This distribution reflects the diverse skill set required for professionals to succeed in the evolving landscape of software engineering and technology. The results highlight that technical domain knowledge, such as database expertise, software engineering principles, and software architectures, remains a fundamental requirement for software professionals, emphasizing the theoretical foundation necessary for specialized roles in the industry.

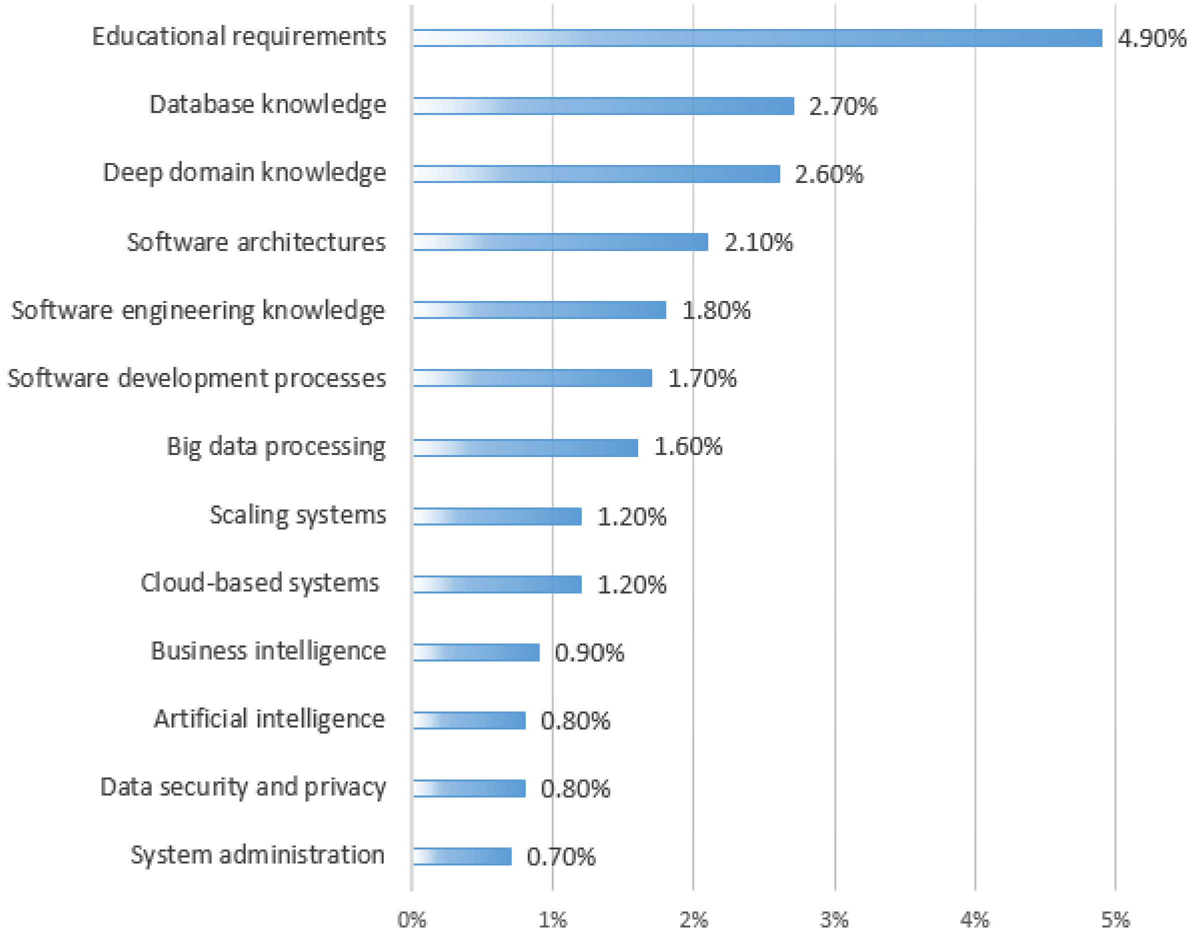

Figure 2. Topics and rates for technical knowledge and skills

The focus on higher education, particularly in areas like computer science and engineering, suggests that formal academic training continues to be highly valued, playing a crucial role in developing essential technical knowledge for software development. The identified trends indicate that technical domain knowledge forms the backbone of software expertise, with critical concepts such as software architectures and engineering methodologies being consistently demanded by employers in the field.

3.2. Software development and programming skills

Of the 52 topics identified in the analysis, 22 were categorized under the broad competency area of "software development and programming skills", which carried a significant 52% weight across the overall distribution. These 22 topics, detailed in Figure 3, are accompanied by their assigned topic names and corresponding percentage values. This breakdown provides a clear understanding of the thematic focus within this competency area, emphasizing its prominence in the software industry. The findings underline a critical emphasis on application development, programming, and coding expertise, as well as the specific knowledge areas required for success in the field of software development. Notably, new-generation scripting languages and leading programming languages such as Java, Python, and JavaScript stand out as the most emphasized qualifications. These findings highlight the importance of proficiency in versatile and widely used technologies, reinforcing their relevance in modern software development environments.

The most prominent topics include "Scripting programming languages" (4.70%), "Javascript frameworks" (4.40%), "Mobile development" (4.20%), "Application development" (4.10%), and "Web services development" (3.80%). These topics represent the core technical competencies demanded by the industry. "Scripting programming languages" and "Javascript frameworks" emphasize modern tools for front-end and back-end development, while "Mobile development" and "Web services development" highlight the growing need for scalable, cross-platform, and user-centric applications. Topics such as "Multi-platform development"

(3.40%), "Object-oriented programming" (3.20%), and "Java programming" (3.10%) reflect the importance of foundational programming principles and creating solutions compatible with diverse platforms.

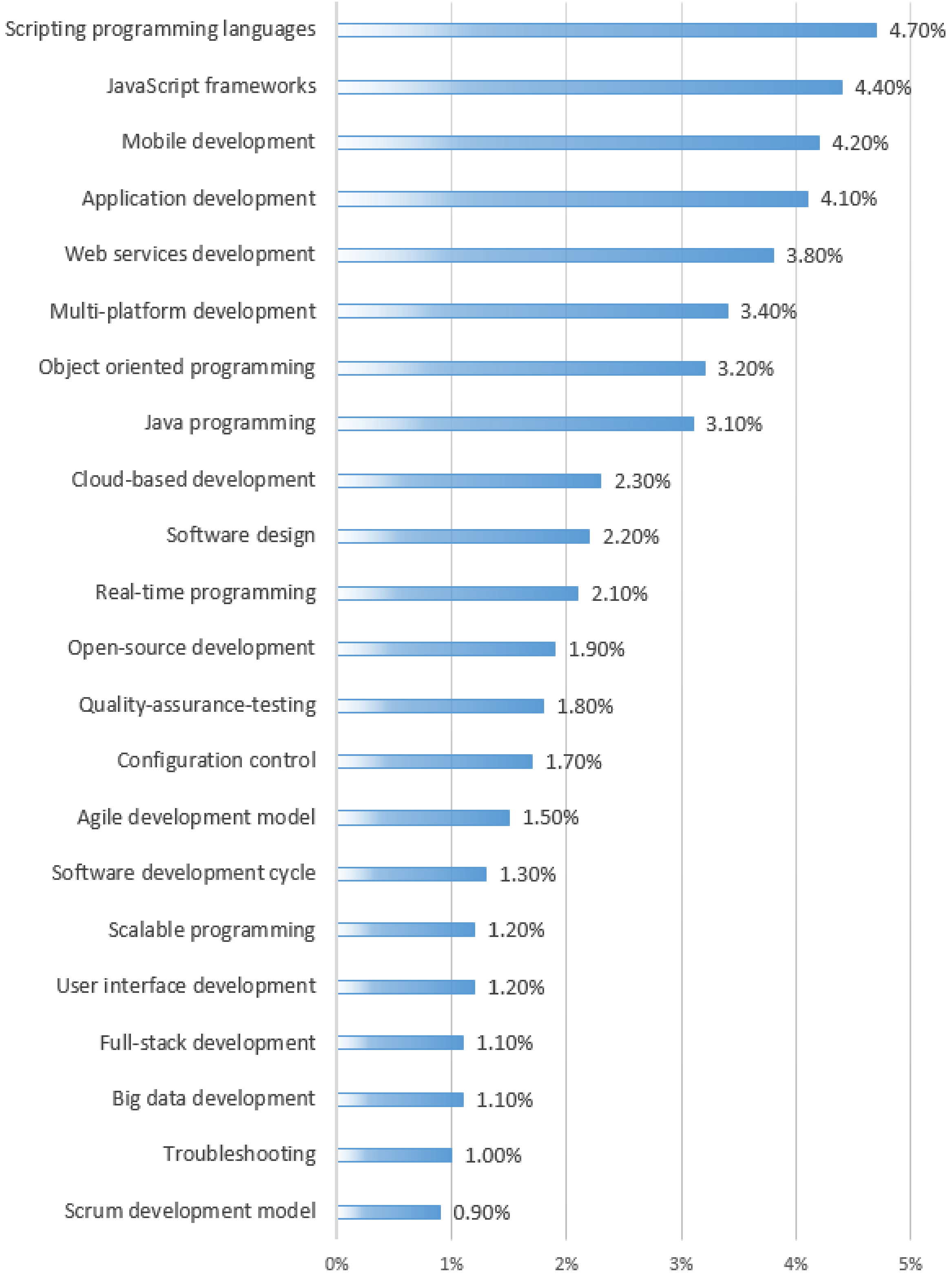

Figure 3. Topics and rates for software development and programming skills

Additionally, "Cloud-based development" (2.30%) and "Software design" (2.20%) emphasize the industry's reliance on cloud architectures and structured software systems. Certain topics, such as "Real-time programming" (2.10%) and "Open-source development" (1.90%), reflect specialized areas of expertise valuable in specific domains. The inclusion of "Quality assurance testing" (1.80%), "Configuration control"

(1.70%), and “Agile development model” (1.50%) highlights the importance of testing, effective project management, and iterative development processes. Emerging areas like “Big data development” (1.10%), “Full-stack development” (1.10%), and “Troubleshooting” (1.00%) indicate growing but less dominant demands for data-driven applications and end-to-end development. Similarly, “Scrum development model” (0.90%), while less frequently mentioned, reflects its ongoing relevance as a widely adopted project management methodology.

This distribution emphasizes a balanced focus on both core programming skills and modern development approaches. The higher weight assigned to topics such as “Scripting programming languages”, “Javascript frameworks”, and “Mobile development” aligns with current trends emphasizing application scalability, user experience, and adaptability to emerging technologies. Additionally, the combination of foundational programming principles (e.g., “Object-oriented programming”) with specialized areas (e.g., “Real-time programming”) reflects the diverse skill set required in software development today. Overall, the topics provides valuable insights into the technical competencies needed for professionals to remain competitive in the dynamic software industry. The findings emphasize that proficiency in programming, paired with an in-depth understanding of contemporary development methodologies, forms the backbone of technical expertise in software development. The ability to utilize scripting languages effectively, develop mobile and web applications, and implement modern programming frameworks is not only a testament to technical competence but also a crucial determinant of professional success in today’s technology-driven landscape.

3.3. Analytical and logical skills

Among the 52 identified topics, ten were observed to fall under the category of “analytical and logical skills.” The total percentage of these topics is 12%. These ten topics, assigned to the analytical and logical skills competency area, are presented in Figure 4, along with their assigned topic names and percentage values. The bar chart in Figure 4 presents the results of a topic analysis conducted on job postings, focusing on analytical and logical skills. The Figure highlights “Requirements analysis” as the most commonly mentioned topic, representing 2.20% of the analyzed postings. This suggests that employers place significant emphasis on understanding and defining project or business requirements, which forms the foundation for effective problem-solving and solution development.

“Data modeling and analysis” follows closely at 1.90%, underscoring the demand for professionals who can structure and interpret complex data sets to inform business decisions. Similarly, “Analytical background” (1.40%) ranks third, reflecting the high value placed on general analytical aptitude across industries. “Machine learning methods” (1.30%), “Problem-solving skills” (1.20%), and “Software metrics” (1.10%) demonstrate the growing importance of technical and quantitative competencies. These skills are particularly relevant in roles that involve leveraging technology and data to optimize processes and outcomes.

On the other hand, “Predictive analytics” (1.00%), while less frequent, still highlights the need for forward-looking insights in decision-making. Lower-ranked topics such as “Decision-making skills” (0.80%), “Business analytics” (0.70%), and “Real-time analytics” (0.40%) suggest a more specialized focus, likely relevant to specific roles or industries. This analysis reveals that job postings for analytical and logical skills prioritize foundational technical abilities, such as requirements gathering and data handling, while also emphasizing the importance of advanced techniques like machine learning and predictive analytics. However, competencies related to real-time decision-making and business-oriented analytics appear less commonly, potentially indicating niche demand or evolving trends in the job market.

These findings underscore the critical role of analytical and logical skills in shaping the modern workforce, particularly in software development processes. Employers prioritize candidates who not only exhibit strong logical reasoning and problem-solving capabilities but also possess expertise in leveraging data-driven tools and methodologies to tackle complex challenges effectively. Key competencies such as requirements analysis and data modeling emerge as foundational elements, reflecting their importance in addressing the multifaceted demands of modern projects. Furthermore, the ability to adapt to evolving technologies and collaborate within multidisciplinary teams has become increasingly essential. As industries continue to embrace digital transformation, the demand for professionals who can combine technical expertise with strategic thinking is likely to grow, further emphasizing the importance of these skills in shaping future career paths.

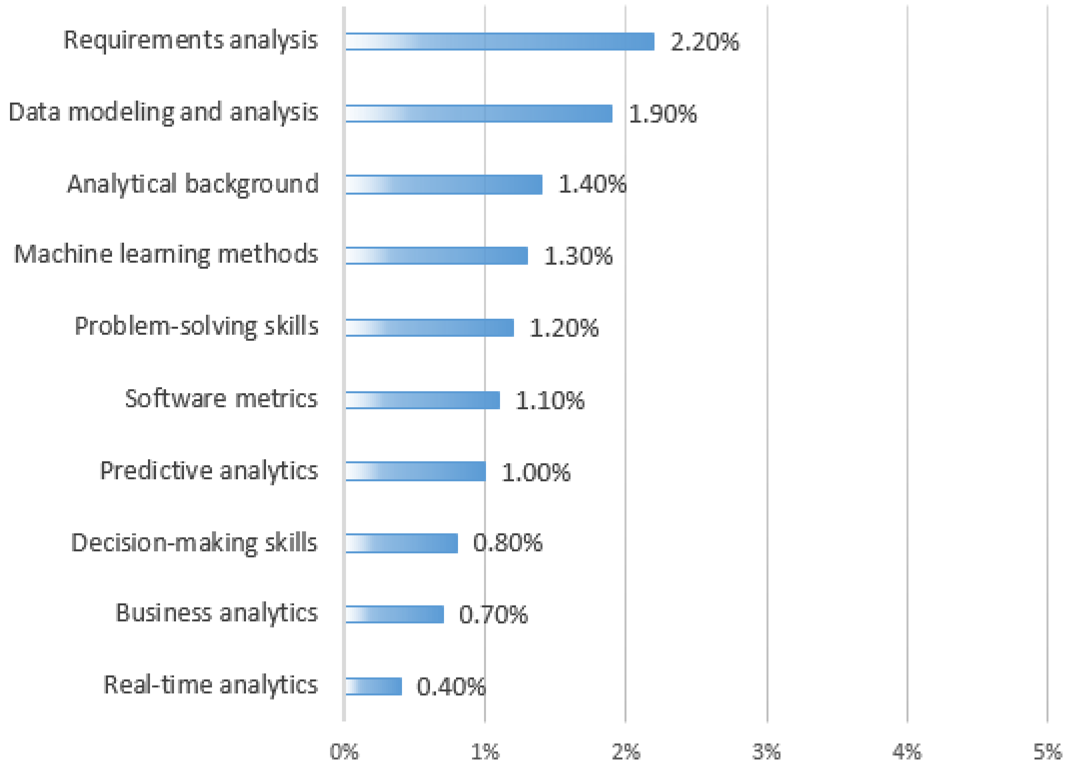

Figure 4. Topics and rates for analytical and logical skills

Additionally, skills like data visualization and various analytical methods are integral to enhancing software development workflows, ensuring precision and efficiency in decision-making. The emphasis on advanced techniques, including machine learning methods and predictive analytics, highlights the evolving nature of expertise required in contemporary software development. This shift demonstrates the growing need for professionals who can seamlessly integrate emerging technologies with traditional analytical competencies to meet the dynamic demands of today's industries.

3.4. Non-technical skills

It was discovered that seven of the 52 topics could possibly be classified as "non-technical skills". These topics make up 13% of the total. Figure 5 lists the seven topics in the non-technical skills competency category, along with the topic names and percentage values. The findings related to non-technical individual skills emphasize various personal competencies that are essential in software development roles, such as communication skills, work experience, teamwork, critical thinking, and creativity.

Figure 5 illustrates the distribution of key interpersonal and non-technical skills identified in job postings. The analysis reveals that "Communication skills" are the most emphasized, accounting for 5.10%, indicating their paramount importance across industries. Effective communication is often critical for collaboration, client interactions, and team coordination, making it a cornerstone of professional success. Following this, "Working experience" ranks second at 2.90%, highlighting employers' preference for candidates with practical, hands-on experience. This suggests that prior experience is highly valued, likely reflecting the importance of applied knowledge and familiarity with workplace dynamics. "Teamwork skills" (1.80%) emerge as the third most mentioned topic, reinforcing the significance of collaborative abilities in achieving organizational goals. This is complemented by "Critical thinking and creativity" (1.10%), which points to a growing demand for innovative problem-solving and strategic thinking in the workplace.

Other notable mentions include "Project management" (0.80%), "Motivation and engagement" (0.70%), and "Organizational skills" (0.60%), indicating that while these skills are less frequently highlighted, they remain essential for certain roles, particularly in leadership and operational contexts. This analysis highlights a holistic approach to skill requirements in job postings, specifically in the software development domain, where interpersonal, creative, and professional competencies intersect to address the rapidly evolving demands of the field. Among these, communication skills are identified as the backbone of effective collaboration in software development. In this highly interdisciplinary domain, developers must articulate complex technical concepts to both technical and non-technical stakeholders, ensuring that requirements are clearly understood and solutions are aligned with organizational goals. Strong communication fosters teamwork, reduces misunderstandings during critical development stages, and enhances project outcomes. Furthermore, in agile and cross-functional development environments, the ability to adapt communication styles to diverse team members and contexts is essential for success.

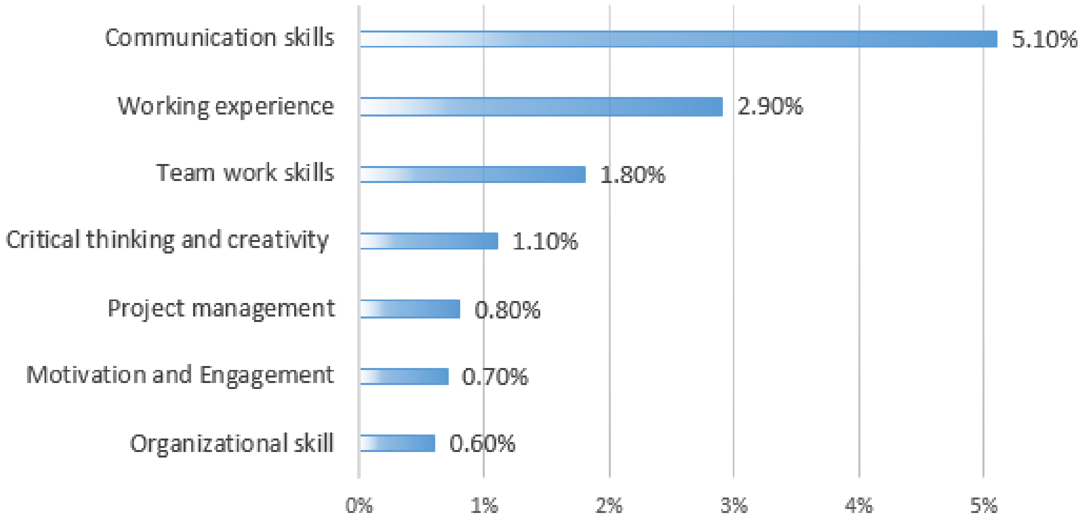

Figure 5. Topics and rates for non-technical skills

Equally critical in software development are critical thinking and creativity, which drive innovation and adaptability. Critical thinking enables software professionals to evaluate problems systematically, analyze vast amounts of data, and identify optimal solutions for highly complex and technical challenges. Whether debugging code, optimizing algorithms, or designing system architectures, critical thinking ensures precision and efficiency, helping teams deliver reliable and scalable software solutions. On the other hand, creativity plays an essential role in fostering innovation within the development lifecycle. It allows developers to think outside conventional frameworks, leading to the creation of groundbreaking software, novel user experiences, and enhanced functionality. Creativity is not limited to front-end or user-facing design; it also drives innovation in areas such as system optimization, automation, and backend efficiency. In an industry characterized by continuous technological advancements, creativity empowers teams to remain competitive by devising unique and forward-thinking solutions.

Together, communication, critical thinking, and creativity form the cornerstone of success in software development. Their interconnected nature ensures that professionals can navigate the multifaceted challenges of the field, effectively collaborate with multidisciplinary teams, and deliver innovative, user-centered solutions. As software development becomes increasingly integral to the success of industries worldwide, these skills are not just desirable—they are indispensable for driving progress, meeting complex project demands, and advancing organizational objectives in this dynamic and competitive landscape.

4. Conclusions

In this study, the rapidly changing dynamics of the software industry and the evolving skill requirements associated with it have been comprehensively addressed. Particularly in recent years, the rapid advancement of technology has transformed the competencies required in software development processes, giving rise to a multidisciplinary framework. Using the LDA-based probabilistic topic modeling method, analysis conducted on a large dataset obtained from online job postings revealed that software developers increasingly require not only technical skills but also soft skills such as analytical thinking, problem-solving, teamwork, communication, and leadership. The findings emphasize the critical role of communication skills, which are foundational in enabling effective collaboration across diverse teams, ensuring clarity in requirements analysis, and facilitating productive exchanges between technical and non-technical stakeholders. Similarly, critical thinking and creativity have emerged as indispensable for tackling the complex challenges of software development, driving innovation, and optimizing solutions. These skills complement technical expertise by empowering developers to analyze problems strategically and approach them with novel perspectives, leading to better software solutions and advancements in system design. This comprehensive skill set, blending technical, programming, analytical, and interpersonal competencies, is crucial for addressing the diverse and complex challenges of the software development field. Both academic institutions and industry leaders should recognize the importance of nurturing this diverse skill set to ensure that developers are well-prepared for the future demands of the profession.

Although the findings of this study provide significant contributions, certain limitations exist. Firstly, the dataset was obtained solely from online job postings, which may not fully reflect the general skill demands of the entire software industry. Additionally, the LDA-based probabilistic topic modeling method may not entirely capture the contextual depth of the topics while revealing latent structures in the texts, potentially overlooking some essential skills. Finally, the dataset used in the analysis pertains to a specific time period, making it challenging to fully capture the continuously evolving trends in software development. Therefore, further studies conducted with more comprehensive and up-to-date data would enhance the validity of the findings. Future research incorporating broader and more updated datasets will enhance the understanding of skill requirements and support educational institutions, employers, and software developers in adapting to the evolving demands of the industry. With further refinements, including expanding the dataset to incorporate additional sources such as LinkedIn, Glassdoor, and Indeed, and enhancing the discussion of specific trends, this work has the potential to become a leading contribution in its field. Additionally, the analysis could be expanded by integrating deep learning-based NLP methods, such as BERT or GPT, to deepen the analysis and enhance its granularity.

This study serves as a critical reference for understanding the current and future skill requirements in the software industry. It offers valuable guidance to software developers in planning their careers by identifying in-demand skills and knowledge areas. Employers can use these findings to better define workforce needs, optimize employee selection processes, and design targeted training programs for skill development. For educational institutions, the study provides actionable insights for updating curricula and developing training programs that align with industry demands, ensuring that graduates are equipped with relevant competencies. Additionally, the findings highlight emerging trends in software development, enabling individuals and organizations to adapt to the evolving landscape of the industry. By addressing the unique needs of these key stakeholders, the study acts as a strategic guide to foster growth, innovation, and alignment with next-generation trends in software development.

Acknowledgement

This study is derived from the doctoral thesis by Fatih Gurcan titled “Determination of Expert Knowledge and Skills for New Generation Software Development Trends Using a Probabilistic Topic Modeling Approach”, conducted in the Department of Computer Engineering at the Institute of Science, Karadeniz Technical University.

Author contribution

Fatih GURCAN – Responsible for conceptualization, methodology design, data collection, formal analysis, writing, reviewing, and editing. Also, he contributed to data interpretation, statistical analysis, task coordination, and validation of the findings. Cemal KOSE – Provided conceptual guidance and acted as an advisor, contributed to software development, data visualization, curation, formal analysis, and assisted with writing, reviewing, and editing.

Declaration of ethical code

The authors of this article declare that the materials and methods used in this study do not require ethics committee approval and/or legal-special permission.

Conflicts of interest

The authors declare that there is no conflict of interest.

References

Aken, A., Litecky, C., Ahmad, A., & Nelson, J. (2010). Mining for computing jobs. *IEEE Software*, *27*(1), 78–85.

Barua, A., Thomas, S. W., & Hassan, A. E. (2014). What are developers talking about? An analysis of topics and trends in Stack Overflow. *Empirical Software Engineering*. https://doi.org/10.1007/s10664-012-9231-y

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*(4–5), 993–1022. https://doi.org/10.1017/9781009218245.012

Burkhardt, S., & Kramer, S. (2019). Decoupling sparsity and smoothness in the dirichlet variational autoencoder topic model. *Journal of Machine Learning Research*, *20*.

Chen, T. H., Thomas, S. W., & Hassan, A. E. (2016). A survey on the use of topic models when mining software repositories. *Empirical Software Engineering*, *21*(5), 1843–1919. https://doi.org/10.1007/s10664-015-9402-8

De Mauro, A., Greco, M., Grimaldi, M., & Ritala, P. (2018). Human resources for Big Data professions: A systematic classification of job roles and required skill sets. *Information Processing and Management*, *54*(5), 807–817. https://doi.org/10.1016/j.ipm.2017.05.004

Debortoli, S., Müller, O., & Vom Brocke, J. (2014). Comparing business intelligence and big data skills: A text mining study using job advertisements. *Business and Information Systems Engineering*, *6*(5), 289–300. https://doi.org/10.1007/s12599-014-0344-2

Egger, R., & Yu, J. (2022). A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts. *Frontiers in Sociology*, *7*. https://doi.org/10.3389/fsoc.2022.886498

Feng, J., Zhang, Z., Ding, C., Rao, Y., Xie, H., & Wang, F. L. (2022). Context reinforced neural topic modeling over short texts. *Information Sciences*, *607*, 79–91. https://doi.org/10.1016/j.ins.2022.05.098

Gurcan, F. (2023a). Identification of mobile development issues using semantic topic modeling of Stack Overflow posts. *PeerJ Computer Science*, *9*, 1–28. https://doi.org/10.7717/peerj-cs.1658

Gurcan, F. (2023b). 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

Gurcan, F., & Cagiltay, N. E. (2019). Big Data Software Engineering: Analysis of Knowledge Domains and Skill Sets Using LDA-Based Topic Modeling. *IEEE Access*, *7*, 82541–82552. https://doi.org/10.1109/ACCESS.2019.2924075

Gurcan, F., & Kose, C. (2017). Analysis of software engineering industry needs and trends: Implications for education. *International Journal of Engineering Education*, *33*(4), 1361–1368.

McCallum, A. K. (2002). MALLET: A Machine Learning for Language Toolkit. In *Http://Mallet.Cs.Umass.Edu*. http://mallet.cs.umass.edu

Moreno, A. M., Sanchez-Segura, M. I., Medina-Dominguez, F., & Carvajal, L. (2012). Balancing software engineering education and industrial needs. *Journal of Systems and Software*, *85*(7), 1607–1620. https://doi.org/10.1016/j.jss.2012.01.060

Stack Overflow. (2024). *Stack Overflow Developer Jobs*. https://stackoverflow.com/jobs

Terblanche, C., & Wongthongtham, P. (2015). Ontology-based employer demand management. *Software: Practice and Experience*. https://doi.org/10.1002/spe.2319

Thomas, S. W., Adams, B., Hassan, A. E., & Blostein, D. (2014). Studying software evolution using topic models. *Science of Computer Programming*, *80*(PART B), 457–479. https://doi.org/10.1016/j.scico.2012.08.003

Vayansky, I., & Kumar, S. A. P. (2020). A review of topic modeling methods. *Information Systems*, *94*. https://doi.org/10.1016/j.is.2020.101582

Wu, X., Nguyen, T., & Luu, A. T. (2024). A survey on neural topic models: methods, applications, and challenges. *Artificial Intelligence Review*, *57*(2). https://doi.org/10.1007/s10462-023-10661-7

Xu, K., Lu, X., Li, Y. fang, Wu, T., Qi, G., Ye, N., Wang, D., & Zhou, Z. (2022). Neural Topic Modeling with Deep Mutual Information Estimation. *Big Data Research*, *30*. https://doi.org/10.1016/j.bdr.2022.100344