A Web Mining-Based Analysis of Job Market Trends on LinkedIn: A Case Study of Sri Lanka

***Abstract*—The dynamic Sri Lankan IT job market necessitates data-driven guidance for stakeholders. This study performs a web mining-based analysis of 380 unique LinkedIn IT job postings advertised between March 2024 and March 2025 to elucidate current employment trends. Employing Python-based web scraping for data collection and comprehensive text analytics including N-gram analysis for skill extraction and sentiment analysis for employer messaging the research examines attributes such as job title, company, location, experience level, and employment type. Key findings highlight dominant demand in Software Development and Network & Systems roles, with a significant concentration of opportunities in the Western Province and within the "IT Services & Consulting" sector. The analysis reveals a strong need for Mid-Senior and Entry-level professionals, predominantly for full-time positions. A balanced profile encompassing technical proficiency, strong soft skills, and relevant educational qualifications (often a computer science degree) is consistently sought. The insights derived aim to inform strategic career planning, targeted workforce development, and evidence-based educational curriculum design.**

Keywords—job market analysis, LinkedIn, web mining, text mining, sentiment analysis, IT industry, Sri Lanka, employment trends, skill demand, data-driven insights

I. INTRODUCTION

The Sri Lankan employment market is undergoing its dramatic changes in terms of economic pressures, technological innovation, and global trends, with the Information Technology (IT) sector being a fast-rising, emerging sector. It is essential that all stakeholders, from IT workers and organizations to tertiary education institutions and policymakers, know the details of this sector in terms of its clear demands of skills, geospatial distribution of opportunities and employers expectations. In the modern business environment, LinkedIn is a leading talent and recruitment platform, offering a granular, real-time dataset for analyzing employer recruitment practices and deducing required employee skills in technology-intensive areas like IT.[1] In this study, we follow a rigorous, evidence-based methodology by systematically scraping IT job listings on LinkedIn with a perspective of gaining a data-driven snapshot of the Sri Lankan IT job market and, therefore, comprehending the prevailing trends of opportunities for employment and mapping out leading employers' expectations.[2] The findings of the web-based study are of high value to IT job applicants by releasing in-demand specializations and skills, for organizations in terms of recruitment benchmarks, for tertiary education institutions in terms of curriculum mapping, and for policymakers in terms of informed decision-making in IT talent, education, and economic planning.

II. LITERATURE REVIEW

Job market research has developed a lot in the last few years from simple interviews to the use of computer software for collecting data especially for a glimpse into Sri Lanka's growing IT job market.[3] Studies show that LinkedIn is extremely well liked by job applicants, with research carried out by finding that 60% of the people prefer using LinkedIn because of its ease of use and also because of the possibility of networking with working individuals, even though only one-third of the users obtain employment via online sites[4]. Sri Lanka's own employment market has its share of issues our country has gone through tough economic periods, showing a deep dip in GDP (-7.8%) and severe problems in the employment market.[4], [5] To add to this, more than 535,000 Sri Lankans moved to work abroad in 2022-2023, radically changing what is left in our local labor pool. In our research, we use web mining (collecting and analyzing data from websites) to the study of job advertisements we break job descriptions into their parts, remove irrelevant words, and study trends in what employers want.[6] This study fills several gaps in what is known about Sri Lanka's IT employment is the gap between official records of employment and what can be observed on sites like LinkedIn,[7] the lack of information on employment in different regions of Sri Lanka,the unsureness of the connection between what is indicated in job advertisements and who companies actually employ, and the need to combine numbers with more perceptive analysis in order to get a realistic impression of how Sri Lanka's IT employment works on LinkedIn..[4], [8]

III. METHODOLOGY

A. Data Collection

LinkedIn was selected as the source of this study for its vast professional job posting pool, which is representative of the Sri Lankan IT job market. Data were obtained through the creation of an in-house Python web scraper, with the requests and BeautifulSoup packages, that fetched job posting information

between 2024 and March 2025 programmatically. In order to circumvent LinkedIn's default result limit of approximately 400 job postings for each search, the scraper was designed to loop through different types of sorting and time frames systematically.

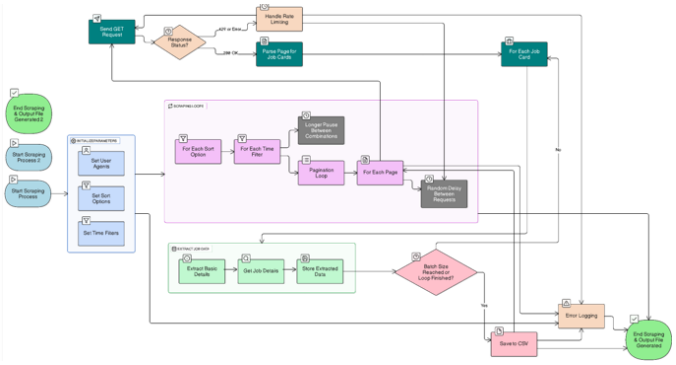

Figure 1 Data Collection workflow

Detection and IP blocking were circumvented during the run of the scraper with the use of rotating user-agents, random pauses between requests, and adaptive throttling when an HTTP 429 status is received. Rich logging functionality was incorporated for progress monitoring and debugging during the course of collecting the data.[7], [9] For each job posting found, the system scraped the full set of attributes that included job ID, job title, company, location, posting date, full description, experience level required, employment type, job function, and industry. All of the web-scraped data were also saved in Comma-Separated Values (CSV), whose choice was due to its high availability of tools for analysis, namely the Pandas library in Python, and therefore allowing easy management of the preprocessing of the data and subsequent statistical analysis.[10]

B. Data Cleaning and Preprocessing

The raw LinkedIn dataset had to be cleaned and pre-processed thoroughly for addressing inconsistencies, null values, and formatting for analysis. Columns with approximately 90% null values (salary, required_skills, applicant_count, company_size, company_industry, job_url, sort_method, and time_filter) were removed in order to preserve analysis integrity, and duplicate job postings were removed on the basis of job_id and description text. For filtering for listings with an IT orientation, a domain-specialized filtering technique was employed utilizing a pre-constructed dictionary of IT-specialty keywords and job titles across a range of different specializations (Software Development, Data Science,[11] Network Systems, Cybersecurity, QA & Testing, IT Support, Project Management, UI/UX Design, Cloud & DevOps, and Technical Writing) in order to programmatically label postings into a new job_category field, from which listings that were not IT-oriented were eliminated from the dataset. Other transformations included/scripting posted_date fields into datetime objects for temporal analysis, extracting province-level details from location fields (with the use of "Unspecified" for general "Sri Lanka" field values for geography-level consistency purposes), performing text normalization activities like lowercase conversion and punctuation stripping for keyword extractibility, as well as removal of the few null values that remained in primary fields (experience_level, employment_type, job_function, industries, and description) in order to have a complete, analysis-ready dataset for Sri Lanka's IT market

C. Data Analysis Techniques

The analysis of the customized IT-focused dataset employed a range of complementary methodologies for extracting complete insights into the IT job market in Sri Lanka. Descriptive statistics enabled the creation of frequency distributions for fundamental categorical variables, including job_category, industries, experience_level, employment_type, and province, thus clarifying basic trends related to IT specializations, hiring industries, experience requirements, and geographical distribution. Temporal trend analysis of monthly and yearly posting methods examined for cyclic or emergent recruitment patterns.[12] Text mining and Natural Language Processing were salient analysis tools,[13] with keyword and N-gram examination of unigrams, bigrams, and trigrams being employed for identifying most-often-referenced skills and qualifications following tokenization and stop word removal, as well as the classification of job benefits and requirements by means of strategic keyword recognition and categorization, followed by sentiment analysis with VADER for analyzing the sentiment in recruitment messages.[14]Comparative and correlational analysis revealed relational patterns, by way of demands for areas of IT specializations and hiring industries, the linkage of experience levels and employment types as heatmaps, skill demands by industries, organizational recruitment preferences, and dynamics of the IT internship environment. The use of the visualization libraries, namely Matplotlib, Seaborn, and Plotly Express, helped in presenting findings in the form of bar charts, pie charts, histograms, heatmaps, and word clouds, and in this manner fully translating complex patterns of data into explanatory visualizations.[15]

D. Ethical Considerations

This study addressed significant ethical aspects of web scraping LinkedIn data. LinkedIn's Terms of Use were respected through the implementation of user-agent rotation and randomized intervals between requests to minimize server burden, with appropriate waiting periods implemented when rate limiting was encountered. All data collection was strictly limited to educational purposes and exclusively gathered from publicly accessible job vacancy listings. The research methodology deliberately avoided the collection of personally identifiable information, focusing instead on general market

patterns rather than individual user behavior. The study-maintained transparency through detailed documentation of collection and analysis processes, with the Python scraping code provided alongside research materials for methodological verification. The collected dataset was stored securely with access restricted to the research team. The findings presented are intended exclusively for providing objective information about Sri Lanka's IT job market without causing harm to individuals or organizations.

IV. Analysis

A. *Overall Landscape of the IT Job Market*

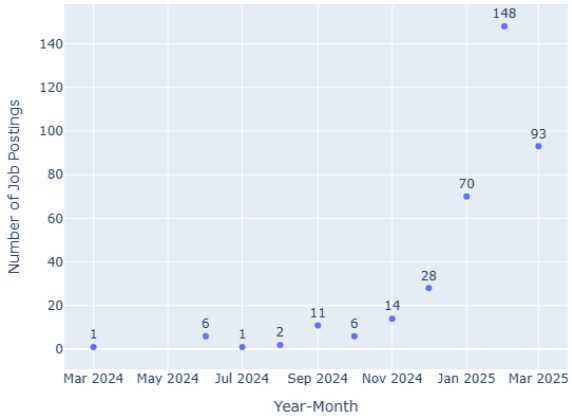

Figure 2 Monthly IT Job Postings in Sri Lanka (Year-over-Year Comparison)

The Sri Lankan IT job openings rise dramatically at certain periods. The number jumps from single digits for the majority of 2024 to 70 openings in January 2025 and a peak of 148 in February 2025. The data also shows Software Development jobs are in highest demand.If you're searching for an IT position in Sri Lanka, strategic application just prior to or during these times of peak activity could significantly improve your chances for securing work.

B. *Demand Hotspots by Industry Sectors and IT Job Functions*

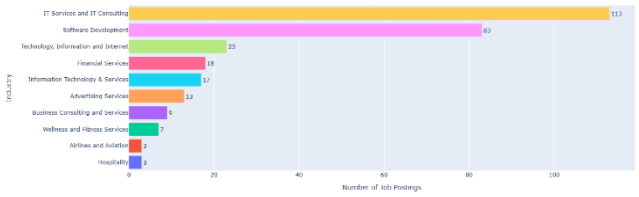

Figure 3 Top Industries Hiring for IT Roles in Sri Lanka

The data strictly shows that the most hiring activities belong to the "IT Services and IT Consulting" sector with 113 out of 380 postings. The "Software Development" sector comes in at second place with 83 postings.

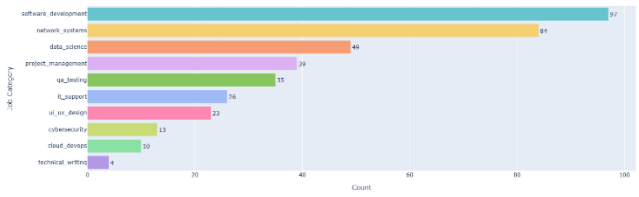

Figure 4 Number of Job Per IT Job Category

Furthermore, full-time jobs are very much available to all levels of experiences, such as high opportunities for entry-level (98 jobs) and mid-senior level (95 jobs) candidates.If you have mid-senior level experience and are prepared to undertake contract work, that's also a viable option, as 34 contract jobs target this experience level.

C. *Profile of IT Roles - Experience Levels and Employment Types*

Figure 3 Distribution of Required Experience Levels for IT Roles

The figures are proof that mid-senior level practitioners are most needed at 34.47% (131 job postings) of positions reviewed. This is unequivocal proof of critical need for experienced IT practitioners who can handle complex tasks and potentially lead groups.

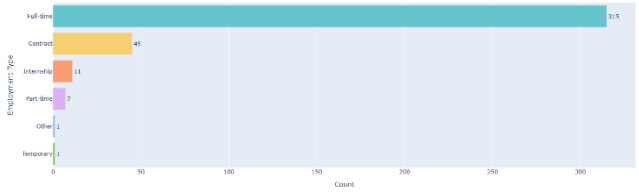

Figure 2 Preference for Emploument Types in the IT Sector

Simultaneously, start jobs make up 28.42% (108 job advertisements), depicting promising opportunities for new entrants to the market. The fact that 82.89% (315 job advertisements) of all positions are full-time jobs is also indicative, as it indicates employers' desire for stable, long-term employment arrangements.

Figure 4 Experience level vs Employment Type

The mid-senior experience level professionals are most likely to be hired, with opportunities remaining high for those of lower levels. The full-time job orientation is consistent with the most typical market offering structure.

D. Geographical Distribution of Job Opportunities

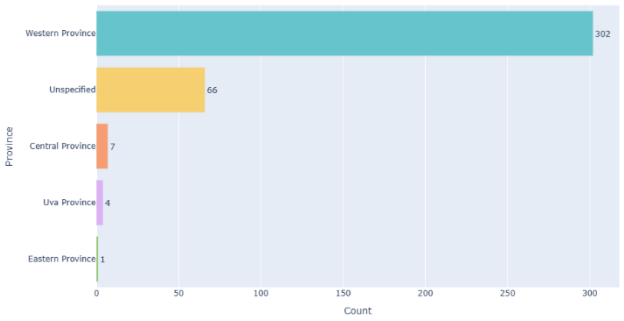

Figure 6 Geographical Distribution of IT Job Postings by Province

The figures clearly demonstrate that the Western Province dominates IT vacancies by a staggering 79.47% (302 job advertisements) of all jobs under analysis. The focus in Colombo is due to the fact that it continues to be Sri Lanka's major economic and technological hub. Also, 17.37% (66 positions) were "Unspecified," most likely indicating remote work opportunities or talent scouting on a country-wide level. Both Western Province and "Unspecified" regions have the most common positions as "Engineering and Information Technology" and generic "Information Technology" positions. If moving or commuting to Colombo is possible for you, this option would optimize your employment opportunities. But also consider applying for jobs with "Unspecified" locations to possibly secure remote work arrangements.

E. Essential Skills and Qualifications in the IT Sector

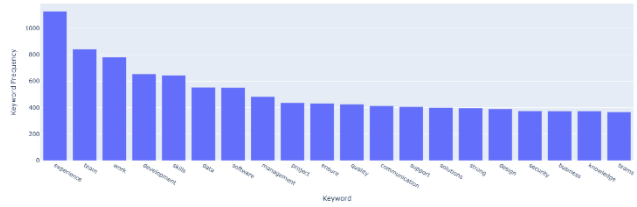

Figure 9 Top 20 Single Word Keywords in IT Job Descriptions

The figure are single-word keyword frequency extraction of words drawn from job titles highlights some of the most common themes. "Experience," "development," "software," "data," and "project" are some of the words that are most commonly repeated, highlighting the technical focus of the profession. "Team," "communication," "skills" (most commonly soft skills), and "management" are some words that are most prominently highlighted, while also highlighting the focus on people skills and team capability. This dual emphasis means that hiring managers seek candidates who possess both solid technical supports and positive soft skills.

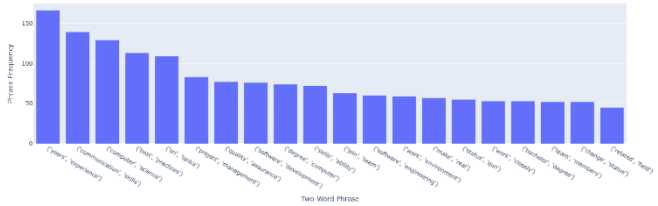

Figure 5 Top 20 Bigram Keywords in IT Job Descriptions

Three-word phrases (trigrams) specify these needs even further. The very common trigram "Degree computer science" highlights the demand for professional study in this field for a great majority of occupations. Phrases like "change status quo" and "get things done" signal the call for individuals being proactive, action-oriented, and perhaps ideally best fit for energetic or innovative company cultures.

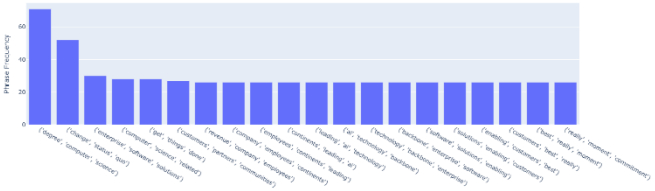

Figure 7 Top 20 Trigram Keywords in IT Job Descriptions

When requirements mentioned in job ads are grouped into broad categories, "Technical Skills" are ubiquitous across the 380 IT job ads analyzed. Next to this, "Experience Level Required" (296 instances), "Soft Skills" (285 instances), and "Educational Requirements" (277 instances) are observed with comparable high frequencies.

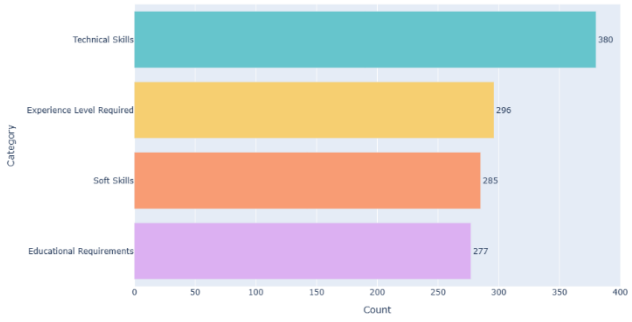

Figure 8 Distribution of Requirement Keyword Categories in IT Job Descriptions

F. Leading Recruiters in the IT Industry

Analysis of IT job postings reveals the top hiring companies in Sri Lanka, led by IFS (32), Dijital Team (22), and YO IT Group (12). Most are software firms or IT consultancies. IFS focused on trainee and PMO roles, emphasizing teamwork and

company culture. Dijital Team offered diverse positions including Cloud Engineers and Security Analysts, often requiring Azure skills. YO IT Group and YO IT CONSULTING highlighted roles involving LLMs, pointing to a demand for AI and Data Science expertise in Python and C++. These insights help job seekers target top employers effectively.

G. Overview of the IT Internship Market

Around 7.11% (27 out of 380) of IT job postings in Sri Lanka were internships, mainly in Software Development and IT Services. This indicates a strong interest in onboarding interns, guiding students' career paths and helping institutions align programs with industry needs

H. Sentiment Expressed in IT Job Descriptions

A VADER sentiment analysis of 380 IT job ads in Sri Lanka showed a high average score of 0.9518, indicating positive language use. Cybersecurity, Technical Writing, and IT Support had the highest sentiment. Data Science roles were slightly lower, reflecting more neutral, technical language..

I. Granular Insights: Industry-Specific and Province-Specific Nuances

A detailed analysis of IT job ads revealed varying skill phrase trends across industries. In "IT Services and Consulting", terms like ‘years experience’ and ‘computer science’ highlight emphasis on qualifications. In contrast, "Software Development" favored phrases like ‘make real’ and ‘enterprise software solutions’, reflecting innovation and systems focus. Regarding benefits, "Professional Development" was most common (97 mentions), followed by "Financial Benefits" (40), "Work-Life Balance" (33), and "Health & Wellbeing" (29). These findings help job seekers, educators, and employers align with industry expectations..

V. RESULTS AND DISCUSSION

A. IT Job Market Trends in Sri Lanka

The trend pattern shows a dramatic rise in vacancies for IT jobs from Sri Lanka, from low single digits in 2024 to 70 in Jan 2025 and an all-time high of 148 in Feb 2025. Early-year recruitment seasonally would be much higher. Software Development remains the leading job category, an indicator of high demand for software solutions from industry.

B. Sector-Wise Hiring Insights

Among 380 job postings analyzed, the IT Services and Consulting sector recorded the highest number of vacancies (113), followed by Software Development (83). These sectors represent the core of digital transformation efforts in Sri Lanka, indicating a robust need for both client-oriented services and custom software development capabilities.

C. Experience Level and Employment Type Distribution

The data highlights a higher demand for mid-senior level professionals, accounting for 34.47% (131 postings). Entry-level roles constitute 28.42% (108 postings), demonstrating a reasonably inclusive job market. Most positions (82.89%) are full-time, emphasizing employers' preference for long-term employment. Contract roles are also present, especially for mid-senior level professionals (34 postings), indicating growing flexibility.

D. Geographical Job Concentration

Job advertisements show a strong concentration in the Western Province, particularly Colombo, with 79.47% (302 postings). An additional 17.37% (66 postings) were unspecified, likely representing remote roles. This suggests that job seekers willing to relocate or explore remote work have better prospects.

E. Skills and Qualifications Analysis

Keyword analysis indicates frequent use of terms such as "experience", "development", "software", "data", and "project", showcasing the emphasis on technical proficiency. Soft skills like "communication" and "team" are also prominent. Trigram analysis emphasized academic qualifications such as "degree computer science" and behavioral traits like "get things done". These findings reflect the dual need for technical capability and soft skill adaptability.

F. Major IT Recruiters in Sri Lanka

Top hiring companies include IFS (32), Dijital Team (22), and YO IT Group (12). IFS focuses on project management and trainee roles, while Dijital Team spans cybersecurity and cloud roles. YO IT Group emphasizes AI and Data Science, including roles requiring skills in LLMs, Python, and C++.

G. Internship Opportunities

Only 7.11% (27 postings) represented internships or trainee positions. These are primarily within Software Development and IT Services, indicating limited but crucial entry-level opportunities for students and new graduates. This highlights a need for improved academic-industry collaboration to enhance practical training access.

H. Sentiment Analysis of Job Descriptions

Using the VADER tool, the average compound sentiment score across job postings was 0.9518, indicating overwhelmingly positive sentiment. Categories such as Cybersecurity (0.9923), Technical Writing (0.9903), and IT Support (0.9888) had the highest positivity. Data Science postings, while still positive, had a lower average (0.8655), possibly due to the objective tone used in technical descriptions.

I. Industry-Specific Insights and Employee Benefits

Phrase frequency varied across sectors. "Years experience" and "best practices" were common in IT Services, while "status quo" and "enterprise software solutions" were more frequent in Software Development. The most commonly offered benefit was Professional Development (97 mentions), followed by Financial Benefits, Work-Life Balance, and Health and

Wellbeing. These patterns show that Sri Lankan IT employers prioritize long-term career growth and employee welfare.

VI. CONCLUSION

This report analyzed 380 LinkedIn IT job postings in Sri Lanka from 2024 to early 2025, revealing key labor market trends. Software Development and Network & Systems roles dominated demand, while growing interest in Data Science and Project Management indicates a shift toward data-driven solutions. N-gram and sentiment analyses showed employers value both technical and interpersonal skills, experience, and formal education—especially in computer science. Most roles targeted Mid-Senior and Entry levels for full-time work, mainly in the Western Province. These findings can guide students, job seekers, institutions, employers, and policymakers to align strategies with the evolving IT job market.

REFERENCES

[1] A. C. Nanayakkara, B. T. G. S. Kumara, and R. M. K. T. Rathnayaka, "A Survey of Finding Trends in Data Mining Techniques for Social Media Analysis," *Sri Lanka Journal of Social Sciences and Humanities*, vol. 1, no. 2, pp. 37–50, Aug. 2021, doi: 10.4038/sljssh.v1i2.36.

[2] G. R. S. R. C. Samaraweera and K. A. N. I. Kariyapperuma, "WHAT DO EMPLOYERS SEEK? DETERMINANTS OF LABOUR DEMAND FOR PROFESSIONAL JOB VACANCIES IN SRI LANKA (WITH SPECIAL REFERENCE TO LINKEDIN NETWORK)." [Online]. Available: https://orcid.org/0000-0001-7103-2702

[3] R. Perera *et al.*, "Spatial analysis of COVID-19 and socio-economic factors in Sri Lanka," Aug. 2021, [Online]. Available: http://arxiv.org/abs/2108.05651

[4] H. Galpaya, L. Senanayake, and S. Perampalam, "Exploring the challenges faced by Sri Lankan workers in web based digital labour platforms 1." [Online]. Available: https://ssrn.com/abstract=3044267

[5] A. Thirulingam and S. Pais, "Discovering the Most Needed Technologies and Skills for IT Jobs on LinkedIn," Te Pukenga, Jul. 2024, pp. 138–149. doi: 10.34074/proc.240118.

[6] B. Shi, J. Yang, F. Guo, and Q. He, "Salience and Market-aware Skill Extraction for Job Targeting," May 2020, [Online]. Available: http://arxiv.org/abs/2005.13094

[7] S. A. C. L. Senarath, S. S. J. Patabendige, and S. Amarathunga, "Sri Lankan Graduate Labour Market: A Status Mismatch," *Journal of Economics, Business and Management*, vol. 5, no. 1, pp. 63–68, 2017, doi: 10.18178/joebm.2017.5.1.487.

[8] A. Rukshan and A. Miroshan, "AN ICT-BASED REAL-TIME SURVEILLANCE SYSTEM FOR CONTROLLING DENGUE EPIDEMIC IN SRI LANKA."

[9] C. Lotfi, S. Srinivasan, M. Ertz, and I. Latrous, "Web Scraping Techniques and Applications: A Literature Review," in *SCRS CONFERENCE PROCEEDINGS ON INTELLIGENT SYSTEMS*, Soft Computing Research Society, 2021, pp. 381–394. doi: 10.52458/978-93-91842-08-6-38.

[10] L. Navaratnam and W. A. Edirisooriya, "A Study of Job Seekers' Perceptions of the Job Portals in Sri Lanka: With Special Reference to 'Millennials' and Generation 'Z,'" *Sri Lanka Journal of Management Studies*, vol. 5, no. 1, Sep. 2023, doi: 10.4038/sljms.v5i1.106.

[11] Y. Alabdulkarim, K. Alruwayti, H. Alsaleh, A. Bablail, and A. Almaslukh, "Analyzing Computing Undergraduate Majors from Job Market Perspective," 2023. [Online]. Available: www.ijeais.org/ijaisr

[12] M. Weerasinghe, K. Weerasinghe, M. Maduranga, and M. Kawya, "Enhancing Web Scraping with Artificial Intelligence: A Review." [Online]. Available: https://www.researchgate.net/publication/379024314

[13] R. Alexander, P. Rukshan, S. Mahesan, S. Lankan Telecom, C. Head Office, and S. Lanka, "Natural Language Web Interface for Database (NLWIDB)," 2013.

[14] J. Thevananth, "LinkedIn as a Gateway to Employment: Analyzing Factors Affecting," *Journal of Information Systems & InformationTechnology (JISIT)*, vol. 9, no. 1, pp. 2478–0677, 2024, doi: 10.13140/RG.2.2.12012.27524.

[15] "Compiled by LIRNEasia for the Asian Development Bank under project number TA-6536 REG: Nowcasting and Disasters: Impact-Based Forecasting and Socioeconomic Monitoring (54113-001)," 2022. [Online]. Available: https://lirneasia.net/futureofwork