Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025 ASEE
Paper ID #48217

Mapping Essential Competencies for Entry-Level Electrical Engineers: A Hybrid NLP and Thematic Analysis Study

Taiwo Raphael Feyijimi, University of Georgia

Taiwo is a highly skilled AI Engineer, Researcher, and Doctoral Student at the University of Georgia who completed his MS in Electrical and Computer Engineering in the College of Engineering. He is currently leveraging AI to tackle simple and longstanding problems in engineering education.

With over a decade of industry experience as a Technology Strategist and Technical Lead, he has established himself as a forward-thinking innovator in AI and EdTech.

His expertise spans Exploratory Data Analysis (EDA), Machine Learning (ML), Natural Language Processing (NLP), and Prompt Engineering Techniques (PETs) with Large Language Models (LLMs). Taiwo is known for his ability to collaborate effectively within and across organizations to meet project goals and drive transformative results. He excels in leading technical teams, offering strategic IT consultations, and implementing solutions that enhance productivity.

VARUN KATHPALIA, University of Georgia

Varun Kathpalia, born and raised in northern part of India, joined EETI as a PhD student in the Spring of 2024. He completed his undergraduate degree in Mechanical Engineering from Chitkara Institute of Engineering and Technology (Punjab Technical University, India) and master's degree in Mechanical Engineering, specializing in Manufacturing & Materials Science Engineering, from the Indian Institute of Technology, Kanpur, India. He has over 4 years of corporate experience with companies such as Hindustan Coca-Cola Beverages Pvt. Ltd. and Saint-Gobain India Pvt. Ltd. (Research & Development). His interest in areas such as improvement in instructional techniques, faculty perspectives and teaching methodologies, drove him towards the domain of Engineering Education. Specifically, the question of how engineering education can be made more effective and engaging fascinated and motivated him to pursue research in this domain.

He is working with his major professor on an NSF funded project dealing with communities and relationships that enable and empower faculty and students in engineering.

Ms. Sarah Jane Bork, University of Georgia

Dr. Sarah Jane (SJ) Bork is an Assistant Professor in Electrical and Computer Engineering with an emphasis on engineering education research. Dr. Bork's research has focused on examining the mental health experiences of engineering graduate students. She has studied different areas (e.g., social factors, engineering culture, etc.) using a variety of research methods (e.g., regression analysis, photovoice, factor analysis, interview data, etc.). Dr. Bork earned her doctorate degree from the University of Michigan's Engineering Education Research Program. Prior to this, she earned both a Bachelor's and Master's degree in Electrical Engineering from The Ohio State University.

Mapping Essential Competencies for Entry-Level Electrical Engineers: A Hybrid NLP and Thematic Analysis Study

Taiwo Feyijimi, University of Georgia

Taiwo Feyijimi is a skilled AI Researcher and Doctoral Candidate in Engineering Education Transformations Institute (EETI) in the College of Engineering at the University of Georgia, Athens, Georgia. He holds an associate degree in Electrical and Electronics Technology and Education, B.Sc. Edu in Physics and Education and MS in Electrical and Computer Engineering. With over a decade-and-a-half of industry experience within tech and education space as a Founder/Co-Founder, EdTech Professional and Advisor to companies, public and private organizations, Taiwo continues to establish himself as a forward-thinking innovator at the nexus of Engineering, AI and Education. His research interests include competency development and leveraging AI tools, technologies and methodologies to enhance ethical research and classroom engagement for advanced problem-solving. Taiwo has developed two pioneering frameworks for integrating AI into qualitative research, which are currently under review for U.S. copyright protection.

Varun Kathpalia, University of Georgia

Varun is a PhD student in Engineering Education Transformations Institute (EETI) in the College of Engineering at the University of Georgia. He completed his undergraduate degree in Mechanical Engineering from Chitkara Institute of Engineering and Technology (Punjab Technical University, India) and master's degree in Mechanical Engineering, specializing in Manufacturing & Materials Science Engineering, from the Indian Institute of Technology, Kanpur, India. He has over 4 years of corporate experience with companies such as Hindustan Coca-Cola Beverages Pvt. Ltd. and Saint-Gobain India Pvt. Ltd. (Research & Development). His interest in areas such as improvement in instructional techniques, faculty perspectives and teaching methodologies, drove him towards the domain of Engineering Education. Specifically, the question of how engineering education can be made more effective and engaging fascinated and motivated him to pursue research in this domain.

Sarah Jane Bork (Assistant Professor), University of Georgia

Sarah Jane (SJ) Bork is an Assistant Professor in the School of Electrical and Computer Engineering (ECE) and the Engineering Education Transformations Institute (EETI) in the College of Engineering at the University of Georgia. Dr. SJ Bork received her B.S. and M.S. in Electrical and Computer Engineering from the Ohio State University in 2017, and her M.S. in Engineering Education Research from the University of Michigan in 2020. As a doctoral candidate in Engineering Education Research at the University of Michigan, Ann Arbor, SJ is studying the mental health experiences of engineering graduate students.

Abstract

Background: Aligning engineering education with the evolving needs of the industry is crucial for producing highly-skilled graduates prepared for the workforce. This study outlines the research methods and some key findings from a broader investigation aimed at identifying the specific competencies that employers seek in entry-level electrical engineering roles, with a particular emphasis on the southeastern United States.

Purpose: This research aims to identify and categorize the essential knowledge, skills, abilities, and dispositions (KSADs) employers look for in this region, providing valuable insights for curriculum development and enhancement strategies to better prepare future graduates.

Methodology/Approach: The study employs a hybrid approach integrating natural language processing (NLP) and thematic analysis. A meticulously curated dataset of 4,585 entry-level electrical engineering job postings from five prominent U.S. job sites ("LinkedIn", "Indeed", "Glassdoor", "CareerBuilder", and "SimplyHired") was analyzed. Machine learning techniques, specifically Latent Dirichlet Allocation (LDA) topic modeling and supervised automatic topic annotation, were utilized to extract and categorize competency-related keywords from the job postings. In parallel, a thematic analysis, grounded in established KSAD frameworks, was conducted to provide a nuanced understanding of the data and capture context-specific insights. This involved manually reviewing the data to understand themes and emerging patterns related to technical expertise, professional skills, dispositions, job roles, and specialized fields.

Results: The analysis identified ten distinct competency themes prevalent in southeastern U.S. entry-level electrical engineering jobs. These themes encompassed a wide array of technical skills (e.g., circuit design, programming, power systems), professional skills (e.g., communication, teamwork, problem-solving), and dispositions (e.g., proactiveness, adaptability, and a commitment to continuous learning). The research provides a detailed mapping of these competencies, revealing their relative importance in the context of the regional job market.

Implications: The findings have substantial implications for electrical engineering curriculum design and teaching practices, providing a data-driven foundation for ensuring alignment with current industry needs in the southeastern United States. The identified KSADs can guide educators in developing targeted courses, workshops, and learning experiences that equip students with the specific skills and attributes sought by employers in the region. Additionally, the study's outcomes can inform career counseling efforts, enabling students to make more informed decisions about specialization and professional development opportunities.

Conclusion: This study underscores the value of integrating NLP and thematic analysis to extract comprehensive competency information from job postings, advancing data-informed practices in engineering education. By providing a detailed analysis of in-demand competencies for entry-level electrical engineering positions in the southeastern U.S., this research empowers educators, policymakers, and industry stakeholders to make informed decisions regarding curriculum development, workforce training, and talent acquisition strategies.

Keywords:

Competency, Electrical Engineering, Computer Engineering, NLP, Machine Learning, Engineering Curriculum, Workplace Readiness.

1. Introduction

In an era marked by rapid technological advancements and shifting industry landscapes, preparing graduates with the skills and knowledge required to meet real-world demands has become a priority in engineering education. Electrical engineering, a discipline central to technological innovation, is experiencing an evolution in workforce expectations, where traditional technical expertise must now be complemented by professional and adaptive skills. This transformation necessitates rethinking how competencies are identified, integrated into curricula, and aligned with the needs of a dynamic job market. Analyzing job postings offers an opportunity to bridge this gap by extracting meaningful insights into the competencies required in the workforce. Leveraging natural language processing (NLP) and machine learning methods, researchers have developed systematic approaches for processing and analyzing unstructured textual data at scale. These tools have been instrumental in identifying trends, patterns, and emerging skills in specific industries, enabling actionable insights for academic institutions and policymakers [1-2]. However, these approaches often face challenges in contextual interpretation and scalability, necessitating hybrid methods that integrate computational rigor with domain expertise.

2. Research Objective and Purpose

This study aims to propose and detail a replicable comprehensive methodological framework for extracting competencies from textual data in electrical engineering job listings. A secondary objective is to present key findings on industry-prioritized competencies for entry-level electrical engineers, categorized into a Knowledge, Skills, Abilities, and Dispositions (KSAD) framework. While this paper minimally discusses the implications of few of these findings, future research will critically examine their impact on engineering education, teaching, and learning, particularly in alignment with employer preferences in the southeastern region for easy comparative analysis with other regions in the U.S. The insights provided aim to inform curriculum development and enhancement strategies to better prepare future graduates.

2.1 Research Question

This methodological study is guided by a central research question: ***What are the top 10 competencies [knowledge, skills, Ability and dispositions (KSAD)] most frequently mentioned in job descriptions for entry-level electrical engineering roles?*** This question underpins the methodological development and aims to identify industry-prioritized competencies critical for aligning academic curricula with workforce demands.

3. Methods

Figure 1 provides an overview our methodological flow.

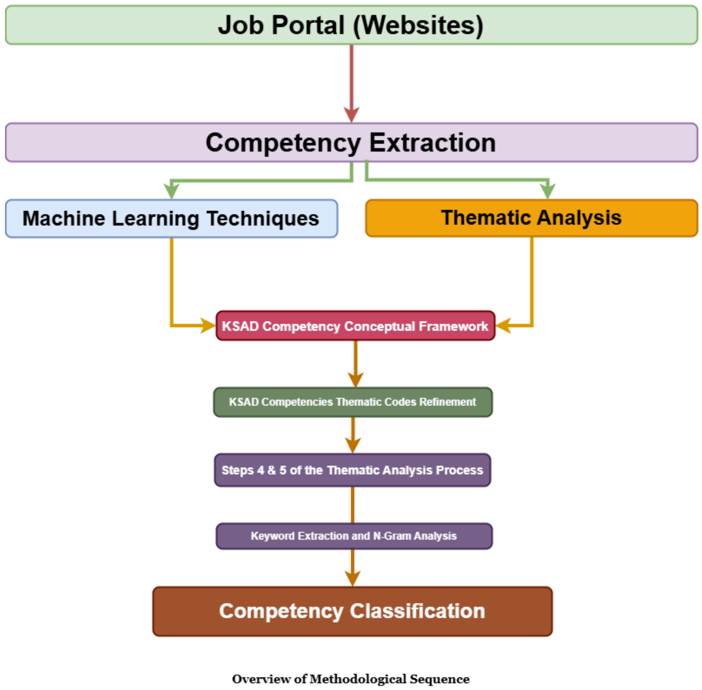

Figure 1. Overview of Methodological Flow

3.1 Positionality Statement

This study examines competencies in undergraduate electrical engineering curricula by analyzing industry job descriptions. The research team, comprising a doctoral researcher and a tenure-track faculty member, brings a blend of expertise in engineering, science, and educational research. These diverse backgrounds shape the team's commitment to addressing gaps between academic preparation and industry demands. The researchers' industry experience highlighted the lack of emphasis on practical skills in education, driving the study's objective to align curricula with job market requirements. A reflective approach acknowledges inherent biases and strives for a balanced, insightful study [3].

3.2 Job Listings Data Sources and Collection

A total of 106,018 electrical engineering job postings were collected from five prominent U.S. job portals: LinkedIn, Indeed, Glassdoor, CareerBuilder, and SimplyHired. These platforms were selected for their broad reach and substantial volume of job advertisements, ensuring a diverse and representative dataset. A custom Python script was developed to automate the extraction of job titles, company names, and job descriptions based on the search parameters "Electrical Engineer" and "Electrical Engineering," identified through a preliminary review of job postings (see schematic flow in Figure 2). The dataset, compiled between May and June 2023, includes detailed job descriptions comprising company overviews and job duties. While the inability to separate these elements during data collection presents a limitation, the dataset serves as a robust resource for analyzing trends in job market demands for entry-level electrical engineers. Despite being collected at a single time point, this dataset enables theoretically grounded and methodologically consistent investigations into workforce requirements and can support standardization in textual job data analysis. This comprehensive dataset forms the foundation for actionable insights into aligning academic programs with industry expectations and meeting evolving labor market demands.

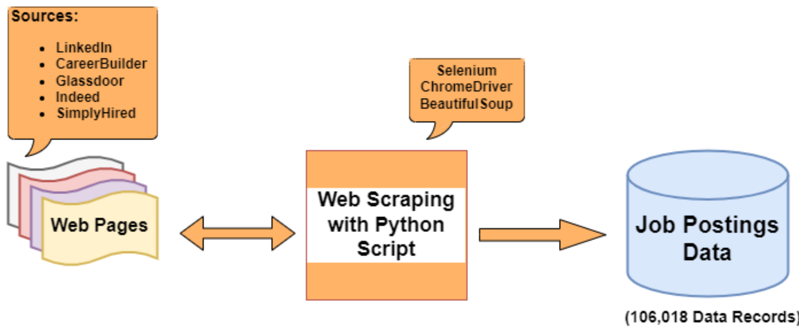

Figure 2: Visualization of Steps Involved in the Data Collection

Table 1. Data Sources for Job Listings

Platform	**Job Listings**
LinkedIn	46,239
CareerBuilder	27,700
Glassdoor	15,964
Indeed	13,209
SimplyHired	2,906
Total	**106,018**

3.3 Data Preprocessing

The preprocessing phase involved the systematic setup of libraries for data handling and text processing, including Pand as for data manipulation and the Natural Language Toolkit (NLTK) for tokenization and stopword removal. Special characters and redundancies were eliminated to ensure data consistency. A manual review of job titles identified discrepancies, leading to title mapping and standardization (see Appendix A1). For example, titles such as "Engineer Development Program - Entry Level Power Engineer" were simplified to "Power Engineer." Custom stopwords were created to address job title-specific noise. To focus exclusively on entry-level electrical engineering (EE) roles, experienced positions were excluded, ensuring the dataset aligned with undergraduate-level competencies. In other words, in constructing our dataset for further analysis, entry-level roles were implicitly defined by exclusion. Specifically, we removed job postings containing keywords indicative of senior positions, including "Senior," "Sr," "VP," "Manager," among others. This was as a result of our knowledge of the dataset during manual review and initial data exploration. This filtering process, detailed in our NLP pipeline, ensured that the resultant dataset comprises positions most suitable for candidates with undergraduate preparation, aligning with the study's focus on entry-level opportunities. Duplicate job postings were removed by cross-referencing job titles. The dataset was further refined by filtering for job postings from 12 southeastern U.S. states, resulting in a final dataset of 4,585 entries. This regional focus enables a targeted analysis of local labor market demands. The processed data was stored in a Pandas DataFrame for subsequent analysis and keyword extraction, establishing a clean and structured foundation for further research.

3.4 Stopword Management

Stopwords, which are common words that typically do not contribute significant meaning to the analysis, are retrieved from NLTK. The standard list of stopwords was extended with additional context-specific words to enhance the precision of the keyword extraction process (see Appendix A2). This tailored stopword list is essential for filtering out irrelevant words and retaining only those with substantive value.

3.5 Data Analysis

3.5.1 Job Title Count

The analysis began with aggregating job titles from the dataset of 4,585 entry-level electrical engineering jobs across 12 southeastern U.S. states. A frequency count using an NLP engine revealed the distribution of roles within the region, offering insights into regional demands. This data was critical for identifying trends and patterns that can inform workforce strategies and curriculum alignment to meet industry needs.

3.5.2 Data Analysis with NLP

The analysis focused on identifying keywords related to “electrical engineering” and “electrical engineer” within the job descriptions. The goal was to extract and categorize competencies, including knowledge, skills, abilities, and dispositions (KSAD). Keywords were divided into hard (technical) and soft (professional) skills to capture the full spectrum of employer expectations. Figure 7 visually summarize the data gathering and thematic analysis processes.

3.5.3 Competency Extraction

Competency identification employed a dual approach, integrating machine learning-based topic modeling and manual thematic analysis. Thematic analysis, following Braun and Clarke's framework [4], ensured that both general trends and nuanced competencies were captured. This hybrid method offered a comprehensive understanding of the required KSAD for electrical engineering roles [2, 5].

3.5.3.1 Machine Learning Techniques

3.5.3.1.1 Topic Modeling with Latent Dirichlet Allocation (LDA)

LDA was employed to categorize competencies in job descriptions by clustering keywords into distinct topics. LDA is a widely used topic modeling technique that effectively analyzes large volumes of uncategorized text by conceptualizing each document as a mixture of topics, with each topic represented as a discrete probability distribution over keywords [6-7]. This probabilistic approach enables LDA to group related words that frequently co-occur, identifying key themes and differentiating meanings in cases of polysemy [8]. As a generative statistical model, LDA replicates the document creation process by associating words with topics, offering a succinct representation of a document's content. This enhanced the demarcation and visualization of skills required for entry-level electrical engineering positions, thereby improving analytical precision. However, LDA treats documents as "bags of words" without capturing relationships between topics, and manual removal of stopwords is necessary to optimize results. While effective for keyword clustering, this limitation highlights the need for complementary approaches to uncover more complex textual relationships.

3.5.3.1.2 Supervised Automatic LDA Topic Annotation

The supervised annotation of LDA topics integrated automated techniques with human judgment and domain expertise to generate meaningful and document-relevant topic labels. The Gensim library in Python was utilized for topic modeling, with model parameters and the number of topics determined through experimentation on a corpus of entry-level electrical engineering job postings [9]. (Testing topic counts ranging from 5 to 40, the Normalized Pointwise Mutual Information (NPMI) score was used to evaluate topic quality, aligning with human assessments of coherence [10]. Empirical analysis identified 10 topics as optimal for the dataset. For each topic, the top 20 keywords with the highest probability were extracted, revealing the distribution of words across topics. A manual review of these keywords, informed by domain knowledge, was conducted to understand underlying themes and refine preliminary labels. This hybrid approach ensured that the annotated topics were both contextually relevant and analytically robust.

3.5.3.1.3 Dimensionality Reduction and Visualization

Dimensionality reduction techniques such as Principal Component Analysis (PCA) and t-SNE (t-Distributed Stochastic Neighbor Embedding) were employed to project high-dimensional data into interpretable 2D or 3D spaces. These methods facilitated the visualization of skill clusters, improving interpretability. Complementary techniques, including t-SNE clustering, word clouds, and graph networks, were used to explore topic interconnectedness and suggest potential themes. Word clouds provided a visual representation of key terms within each topic, while graph networks highlighted relationships between topics. Refinements to topic labels were guided by topic coherence metrics, ensuring consistency and logical alignment across the dataset [11]. Annotated topics were subsequently used to label documents and text segments, enabling a thematic mapping analysis that captured the interconnected structure of topics. This comprehensive approach ensured that visualizations and annotations were both analytically robust and contextually meaningful.

3.5.4 Thematic Analysis Method

3.5.4.1 Steps 1–3: Understanding the Data, Creating Initial Codes, and Searching for Themes

The thematic analysis process (Figure 3) began with familiarization during the data cleaning and processing phase. Initial codes were created to uncover prominent patterns and emerging industry trends within the large dataset of electrical engineering (EE) job postings. A manual approach was employed to iteratively review raw job postings, immersing the analysis in the industry context and identifying core areas of interest, such as technical expertise (e.g., circuit design, power systems), professional skills (e.g., communication, leadership), dispositions (e.g., innovation, flexibility), specific roles (e.g., systems engineer, RF engineer), and specialized fields (e.g., renewable energy, automation) [12]. Systematic evaluations revealed variations in required skills and qualifications across postings, shedding light on unique job requirements and emerging trends in the EE sector. The analysis followed a deductive approach, leveraging research questions to guide the identification and refinement of themes, while revisiting foundational knowledge of the EE field to enhance interpretive depth [4, 13]. This preparatory phase was critical for generating effective codes and themes, ensuring a thorough understanding of interconnected industry concepts and requirements. It should be noted however, that the thematic analysis was conducted on a few 900 sample entries of the job descriptions, approximately 20%, to extract key competencies.

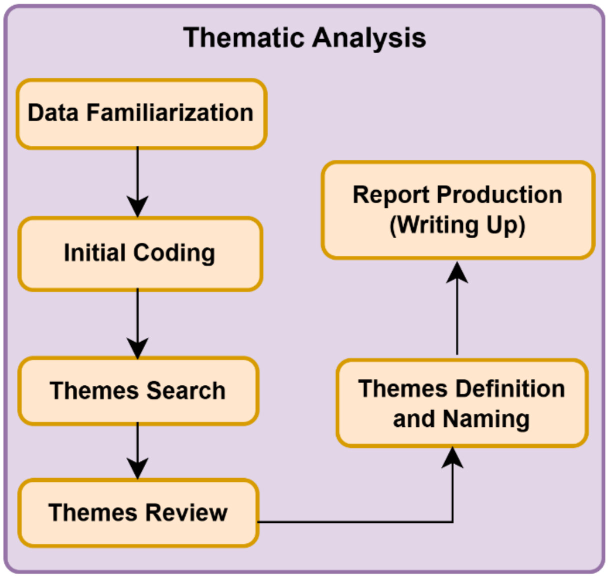

Figure 3. Thematic Analysis Process following Braun & Clarke (2006) Model

3.5.4.2 KSAD Competency Conceptual Framework

The competency framework was developed through a manual review of recurring keywords and phrases, supported by advanced tools such as graph networks and word clouds. These methods enabled the identification of interconnected topics and the categorization of competencies into four main areas: **knowledge, skills, abilities, and dispositions (KSAD),** as given in Figure 4. This categorization provided a robust theoretical basis for competency classification [4, 14].

- **Knowledge:** Represents a theoretical and practical understanding, including core principles, methodologies, and advancements in electrical engineering. It focuses on technical areas such as electrical systems, engineering principles, and scientific concepts fundamental to the discipline [15].
- **Skills:** Divided into:
  - **Hard Skills:** Tangible, teachable technical abilities (e.g., programming, circuit design, and data analysis) [16].
  - **Soft Skills:** Interpersonal and professional competencies (e.g., communication, teamwork, and problem-solving) essential for workplace collaboration and leadership [17].
- **Abilities** refer to innate or developed capacities, including cognitive capabilities like analytical thinking, creativity, and the ability to adapt quickly to new technologies. These are vital for addressing challenges and applying skills effectively across contexts [18].
- **Dispositions (Attitudes):** Behavioral traits and mindsets (e.g., resilience, ethical responsibility) that influence professional engagement and development [19].

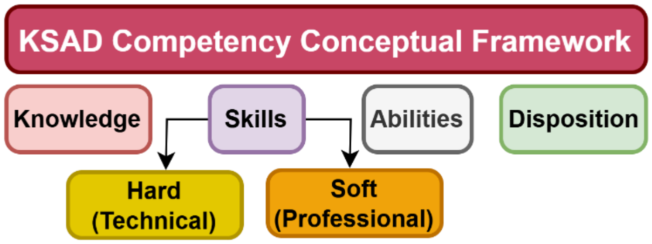

Figure 4. KSAD Competency Conceptual Framework

3.5.5 Methodological Framework for Keyword Extraction and Forward-Backward N-Gram Analysis

3.5.5.1 Steps 4 and 5: Reviewing, Naming and Defining Themes

This framework outlines a systematic approach for extracting keywords and conducting forward-backward n-gram analysis on job description text, emphasizing steps 4 (reviewing and naming themes) and 5 (defining themes) of the thematic analysis process. The methodology integrates text preprocessing, keyword extraction, n-gram generation, and skill classification to build a structured dataset suitable for further exploration. In this process, codes serve as descriptive markers for data segments, summarizing key features, concepts, or ideas. When applied to job postings for electrical engineering roles, these codes highlight recurring competencies, qualifications, and job requirements, providing a foundation for identifying broader themes during subsequent analysis. The framework ensures a robust and meaningful examination of textual data, facilitating the classification and interpretation of KSAD competencies.

3.5.5.2 Data Visualization and Preliminary Coding

Initial data visualization (Figures 8 & 9, in Appendices B & C, respectively) facilitated the identification of salient attributes within the dataset, enabling a focused examination of key features. Preliminary labels were assigned to these attributes, forming the foundation of a structured coding strategy. The choice of coding approach was carefully considered, balancing inductive coding – which identifies patterns emerging from the data – and deductive coding, guided by our research questions and classification framework. This dual approach ensured a robust and methodologically sound coding process tailored to the dataset's complexity and context.

3.5.5.3 Reviewing Job Postings: Identifying Key Insights

The review of job postings focused on extracting text segments that highlighted critical competencies and qualifications for entry-level electrical engineering roles. For example, job descriptions often specified requirements such as "strong fundamentals in circuit design" and "working knowledge of PCB layout software," underscoring the importance of proficiency in electrical circuitry and experience with Printed Circuit Board (PCB) design. This meticulous analysis provided precise insights into the technical skills prioritized by employers.

3.5.5.4 Labeling and Coding Strategy

Descriptive labels were assigned to text segments to systematically capture their essence within the analytical framework. Labels such as "K" (Knowledge), "SS" (Soft Skills), and "HS" (Hard Skills) were employed to categorize competencies accurately. Consistency was ensured through periodic revisions, maintaining the accuracy and relevance of coding as the dataset evolved. Metrics like Frequency, Pointwise Mutual Information (PMI), and Cosine Similarity supported the coding process, with detailed examples provided in Table 2. This strategy facilitated a robust and structured approach to competency classification.

Table 2. Illustrative Bigrams Juxtaposed with Frequency, PMI Score, and Cosine Similarity Post-coding

Bigram	Freq	PMI	Cosine Similarity	Percentage (%)	Knowledge (K)	Skills	Ability (A)	Disposition (D)
electrical engineering	3547	2.9	0.85	0.26	K			
year experience	2243	3.45	0.64	0.16				
control system	1734	3.09	0.75	0.13	K			
sexual orientation	1690	6.67	0.84	0.12				
equal opportunity	1635	5.11	0.85	0.12				
gender identity	1529	6.38	0.84	0.11				
degree electrical	1493	3.27	0.72	0.11	K			
electrical engineer	1429	2.74	0.74	0.1	K			
veteran status	1370	5.32	0.45	0.1				
race color	1345	7.21	0.58	0.1				
orientation gender	1276	6.18	0.83	0.09				
opportunity employer	1266	4.91	0.8	0.09				
national origin	1214	6.25	0.83	0.09				
color religion	1179	6.31	0.76	0.09				
electrical system	1132	1.86	0.69	0.08	K			
regard race	1082	6.07	0.63	0.08				
communication skill	1039	4.19	0.83	0.08		SS		
protect veteran	891	5.33	0.9	0.07				
employment regard	876	6.18	0.8	0.06				
receive consideration	873	6.43	0.9	0.07				
religion sex	869	6.17	0.87	0.07				
consideration employment	854	5.6	0.61	0.07				

reasonable accommodation	842	6.85	0.96	0.06				
skill ability	817	3.53	0.74	0.06			A	D
applicant receive	811	5.44	0.87	0.06				
bachelor degree	789	3.13	0.68	0.06	K			
project management	767	4.39	0.8	0.06		HS		
team member	767	4.39	0.8	0.06		SS		D
qualified applicant	748	5.48	0.81	0.06				
include limited	762	4.7	0.73	0.06				
engineering relate	731	2.69	0.69	0.06		HS		
employment opportunity	721	3.97	0.76	0.06				
ability work	709	2.26	0.85	0.05			A	D
work environment	708	3.49	0.8	0.05		HS		
dental vision	702	6.5	0.72	0.05				
system design	697	1.47	0.69	0.05		HS		
power system	697	2.62	0.87	0.05		HS		

3.5.5.5 Sentence Reconstruction

Sentence reconstruction was employed to address the contextual ambiguity of certain bigrams encountered during analysis. This process involved piecing together surrounding text to derive accurate interpretations, ensuring correct categorization and refining coding integrity. By integrating both syntactic and semantic associations, sentence reconstruction provided essential context to fragmented information, significantly enhancing coding precision and reliability. For example, the bigram "without regard," as depicted in Figure 5, lacked clarity until reconstructed within the context of Diversity, Equity, and Inclusion (DEI). Similarly, Figure 6 illustrates how this methodology clarified ambiguous bigrams, enabling more nuanced coding. This approach substantially improved the validity and quality of the study, offering a robust framework for analyzing job postings in the electrical engineering field and yielding valuable insights into industry requirements.

1. All qualified applicants will receive consideration for employment without regard to age, race, color, religion, sex, sexual orientation, gender identity, national origin, disability or protected veteran status.us.
2. All qualified applicants will receive consideration for employment without regard to race, color, religion, sex, age, national origin or ancestry
3. provides equal employment opportunities (EEO) to all employees and applicants for employment without regard to race, color, gender, religion, sex, sexual orientation, ethnicity or national origin, age, disability, marital status, genetics, pregnancy, or any other protected characteristic as outlined by federal law.
4. Employment decisions are made without regard to race, color, religion, national or ethnic origin, sex, sexual orientation, gender identity or expression, age, disability, protected veteran status or other characteristics protected by law.

Figure 5: Sentence Reconstruction Demonstrated for the Bigram "without regard"

1. 1+ year industry experience solar andor commercial electrical1+ year experience in Operations, LOTO with medium to high voltageIntermediate skills working with MS OfficeWillingness to
2. All other duties as assignedEducation and Experience:Bachelor's degree in Engineering, IT, or in related fieldsYears of Experience: 2+ year experience in industry and related equipmentEnglish
3. Performs other duties as required.QualificationsBasic Job RequirementsBachelor of Science degree in Electrical Engineering, Computer Engineering, or related field with a minimum of three (3) year experience in SCADA, servers and networks, or automation and PLC.Demonstrable proficiency
4. Helper Electricians:Minimum of 2 year experience in electrical fieldKnowledge of conduit bending a plus
5. effectively with the Community Manager all maintenance issues. Requirements1 year experience as a residential maintenance technicianValid
6. relevant industry work experience with no less than 1-year experience in an SAP and/or SCI environment within the past

Figure 6: Sentence Reconstruction Demonstrated for the Bigram "year experience"

3.5.5.6 Keyword Extraction and N-Gram Generation

The keyword extraction process began with tokenizing the input text and converting it to lowercase for consistency. Stopwords were removed, and non-alphabetic characters were filtered out to ensure textual clarity. This refinement yielded a unique set of keywords, which served as the foundational dataset for subsequent n-gram analysis, enabling precise identification of patterns and associations in the text. The methodology advanced to generating forward and backward n-grams, including bigrams, trigrams, and quadgrams. Each n-gram was constructed by analyzing the surrounding context of identified keywords. For instance, bigrams paired a keyword with its immediately preceding word, trigrams included two preceding words, and quadgrams encompassed three preceding words. This approach provided a detailed understanding of keyword usage patterns and associations, enabling deeper contextual insights.

3.5.5.7 Assessing Word Combinations via Advanced N-Gram Analysis

The analysis of word combinations extended beyond bigrams to include trigrams and quadgrams, validating initial findings with three key metrics: frequency, Pointwise Mutual Information (PMI) score, and cosine similarity as detailed in Table 3.

- **Frequency** identified recurring n-grams to uncover prominent themes in the dataset.
- **PMI Score** quantified the likelihood of word co-occurrence beyond random chance, filtering out weak associations.
- **Cosine Similarity** measured semantic alignment of word vectors in a multi-dimensional space, ensuring refined n-gram selection.

Benchmark values were established to optimize analysis:

- A **PMI score** baseline of 3 ensured meaningful co-occurrences.
- A **frequency threshold** of 600 emphasized stronger term relationships.
- A **cosine similarity threshold** of 0.5 confirmed moderate to high semantic coherence.

Table 3: Key Bigram Associations in Text Data Analysis

Bigram	**Frequency**	**PMI**	**Cosine Similarity**	**Percentage**
electrical engineering	3547	2.9	0.86	0.26
year experience	2243	3.45	0.65	0.16
control system	1734	3.09	0.76	0.13
sexual orientation	1690	6.67	0.85	0.12
equal opportunity	1635	5.11	0.86	0.12

Table 3 is a representative sample of the key bigram associations in our dataset. It presents a set of bigrams (two-word phrases) along with their frequency, PMI, Cosine Similarity, and Percentage. "Electrical engineering" appears most frequently (3,547 times) and has a strong cosine similarity (0.86), indicating a well-established contextual association. "Sexual orientation" and "equal opportunity" have the highest PMI values (6.67 and 5.11, respectively), suggesting these terms appear together more often than by random chance, indicating a strong semantic relationship. The varied cosine similarity values highlight different levels of contextual cohesion among the bigrams.

3.5.5.8 Skill Classification

Quadgrams were classified into technical skills, professional skills, and dispositions using tailored regular expressions. This step ensured that extracted competencies were aligned with predefined categories, capturing both explicit and implicit skill requirements.

3.5.5.9 Data Annotation and Visualization

Annotated datasets incorporated classification results, enabling subsequent analyses. Visualizations, including bar charts and word clouds, highlighted the distribution of competencies, providing actionable insights for workforce alignment.

3.5.6 Ensuring Reliability and Validity

3.5.6.1 Credibility and Peer Debriefing

Credibility was ensured through peer debriefing sessions, incorporating feedback from external reviewers to refine coding processes. This approach validated the analytical rigor of the framework [20].

3.5.6.2 Prolonged Engagement and Documentation

Continuous engagement with the dataset and meticulous documentation enhanced trustworthiness, aligning with best practices for qualitative research [21].

3.6 Contributions to Engineering Education Research (EER)

- **Innovative Methodology:** Combining NLP and thematic analysis fosters a replicable model for competency analysis in diverse domains.
- **Practical Implications:** Results provide actionable insights for curriculum alignment with real-time job market demands.
- **Scalability:** The framework's adaptability extends its utility to other engineering disciplines.

3.7 Limitations and Future Work

Limitations include reliance on static datasets and challenges in contextual ambiguity. Future research should incorporate dynamic, real-time analyses and advanced semantic models for improved generalizability.

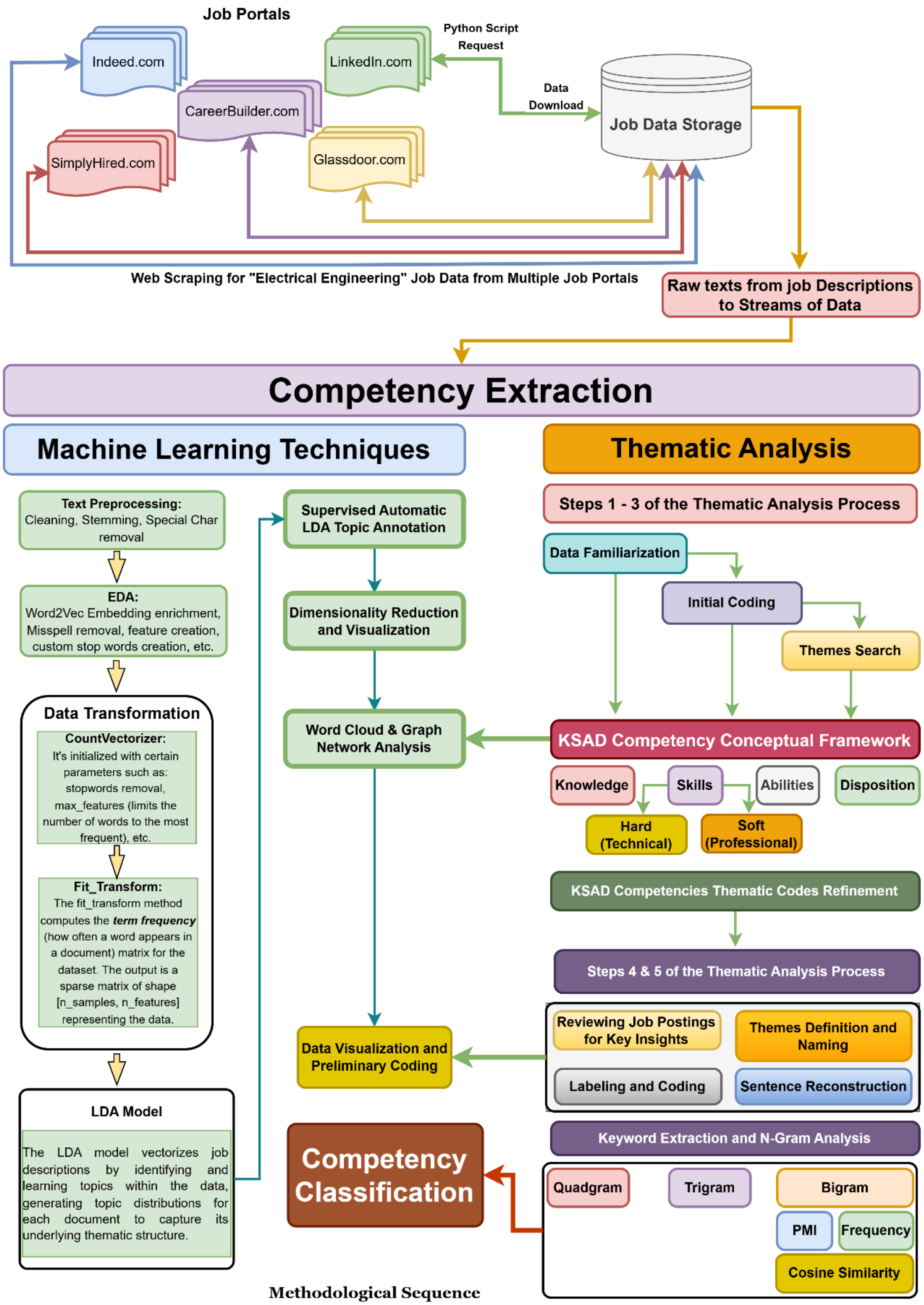

Figure 7. A visual representation of our comprehensive and novel methodology presented in a sequential workflow

4. Results: Analysis of Key Competencies for Electrical Engineering Roles

The modern electrical engineering (EE) landscape is characterized by a diverse array of competencies that extend beyond traditional technical skills. An analysis of job descriptions reveals a hierarchy of hard skills, soft skills, and dispositions that are essential for success in entry-level positions (detailed in Table 4). This analysis, grounded in the Knowledge, Skills, Abilities, and Dispositions (KSAD) framework, underscores the multifaceted nature of contemporary EE roles.

Table 4. Top Hard, Soft and Dispositional Competencies from EE Job Descriptions

Hard Skills	Soft Skills	Disposition
PLC (Programmable Logic Controller) (10,637 mentions)	Leadership (500 mentions)	Innovative (372 mentions)
Automation (7,456 mentions)	Communication Skills (321 mentions)	Project Management (171 mentions)
Control Systems (4,562 mentions)	Initiative (293 mentions)	Analytical (148 mentions)
Instrumentation (5,105 mentions)	Collaboration (229 mentions)	Flexibility (103 mentions)
Simulation (3,818 mentions)	Presentation Skills (204 mentions)	Attention to Detail (97 mentions)
SCADA (3,311 mentions)	Problem-Solving (183 mentions)	Timely Manner (76 mentions)
Sensors (3,312 mentions)	Teamwork (163 mentions)	Self-Starter (69 mentions)
Python Programming (2,735 mentions)	Work Independently (119 mentions)	Interpersonal Skills (69 mentions)
Power Systems (2,739 mentions)	Reliability (143 mentions)	Computer Skills (65 mentions)
MATLAB (1,453 mentions)	Creativity (96 mentions)	Mathematical (51 mentions)

4.1 Prominent Competencies

4.1.1 Hard Skills

PLC (Programmable Logic Controller) programming is the most frequently mentioned hard skill (10,637 mentions), highlighting the demand for expertise in industry automation. Automation (7,456 mentions) and control systems (4,562 mentions) are also highly valued, reflecting the increasing complexity of EE systems. Instrumentation (5,105 mentions) and simulation (3,818 mentions) are critical for precision and efficiency in engineering roles. SCADA (Supervisory Control and Data Acquisition) systems (3,311 mentions), sensors (3,312 mentions), Python programming (2,735 mentions), and power systems (2,739 mentions) indicate the growing integration of software, automation, and energy solutions. MATLAB (1,453 mentions) proficiency remains foundational, balancing established principles with innovative trends.

4.1.2 Soft Skills

Leadership (500 mentions) and communication skills (321 mentions) are highly prioritized, reflecting the need for engineers to manage teams and articulate technical concepts effectively. Initiative (293 mentions), collaboration (229 mentions), presentation skills (204 mentions), and problem-solving (183 mentions) are also crucial for navigating team dynamics and conveying solutions. Teamwork (163 mentions), working independently (119 mentions), and reliability (143 mentions) highlight the importance of interpersonal and emotional intelligence competencies. Creativity (96 mentions) demonstrates the value of innovation in modern engineering roles.

4.1.3 Abilities and Dispositions

Innovative (372 mentions) and project management (171 mentions) dispositions are highly valued for fostering creativity and problem-solving. Analytical skills (148 mentions) and flexibility (103 mentions) emphasize the need for adaptability and forward-thinking in engineering roles. Attention to detail (97 mentions), timely manner (76 mentions), and self-starter (69 mentions) dispositions ensure reliability and compliance in complex tasks. Interpersonal skills (69 mentions), computer skills (65 mentions), and mathematical abilities (51 mentions) are essential for collaboration and teamwork.

5. Discussion

5.1 Competencies in Electrical Engineering Roles: Aligning Education with Industry Demands

Entry-level electrical engineering roles necessitate a comprehensive skill set encompassing technical, professional, and dispositional competencies. Technical skills such as PLC programming, automation, and control technologies are frequently emphasized, reflecting the increasing demand for expertise in industry automation. Complementing these are professional skills like communication, collaboration, and leadership, which underscore the importance of interpersonal abilities in fostering teamwork and problem-solving. Dispositional traits such as innovation, analytical skills, and flexibility highlight the value of adaptability and forward-thinking in engineering roles. These competencies collectively demonstrate the multifaceted demands of modern engineering, where technical expertise is enhanced by strong interpersonal and cognitive capabilities. The discussion of these competencies is guided by the Knowledge, Skills, Abilities, and Dispositions (KSAD) framework.

5.1.1 Foundational Knowledge

Foundational knowledge in electrical engineering centers on essential competencies such as electrical theory, programming, and digital literacy, including CAD platforms and BIM tools like REVIT, which support critical tasks like circuit design and system analysis. Emerging areas such as embedded systems, IoT, SCADA systems, and predictive modeling tools like LabView and FPGA Design reflect the field's shift toward smart systems and real-time control technologies. Despite these advancements, traditional skills like power systems analysis and MATLAB proficiency remain foundational, demonstrating the discipline’s balance between established principles and innovative trends. The integration of automation, IoT, and data analytics into EE highlights the demand for engineers who can blend theoretical expertise with practical, interdisciplinary applications.

5.1.2 Essential Skills

5.1.2.1 Hard Skills

Hard skills are the cornerstone of Electrical Engineering, encompassing the technical expertise acquired through formal education, practical training, and professional experience. Prominent skills such as PLC programming, automation, control systems, instrumentation, and simulation are highly sought after, reflecting the industry's focus on precision and efficiency. Emerging trends reveal the growing importance of SCADA systems, Python programming, and power systems expertise, highlighting the increasing integration of software, automation, and energy solutions. This interdisciplinary evolution underscores the complexity and dynamic nature of modern EE.

5.1.2.2 Soft Skills

Soft skills are equally critical, with leadership and communication being highly valued (see Table 4). These skills are consistently prioritized, reflecting the need for engineers to manage teams, oversee projects, and effectively articulate technical concepts. Employers emphasize these skills because they enable professionals to navigate team dynamics and convey complex solutions to diverse audiences. The spectrum of soft skills extends to initiative, collaboration, mentoring, and empathy, illustrating the multifaceted demands of modern engineering roles. These interpersonal and emotional intelligence competencies are indispensable in technical environments.

5.1.3 Abilities and Dispositions

Abilities in EE involve the application of both hard and soft skills to real-world engineering challenges. They encompass the capacity to analyze complex systems, synthesize innovative solutions, and adapt to evolving technologies. This intersection of knowledge, technical acumen, and interpersonal competence positions EE professionals to excel in dynamic industry roles. Key dispositions such as innovation, ethical conduct, and analytical abilities are increasingly valued for fostering creativity and problem-solving. Project management and analytical skills highlight the interdisciplinary nature of EE, while adaptability, precision, and ethical standards ensure reliability and compliance. Proactive traits like leadership and self-motivation, combined with strong interpersonal and communication skills, are essential for collaboration and teamwork. Balancing technical competencies such as SCADA and control systems with collaborative traits supports the smooth execution of modern engineering projects. Overall, EE professionals must harmonize technical challenges with collaboration and innovation.

6. Recommendations

Based on these findings, several recommendations can be made for curriculum development, faculty training, and accreditation:

6.1 Curriculum Development:

- **Integrate emerging technologies:** Incorporate modules on IoT, SCADA systems, predictive modeling, and data analytics to reflect the increasing integration of these technologies in EE.
- **Balance traditional and modern skills:** Maintain a focus on foundational skills such as power systems analysis and MATLAB proficiency while introducing newer programming languages like Python and automation tools.
- **Emphasize interdisciplinary learning:** Promote courses that combine electrical engineering principles with software, automation, and energy solutions to foster a holistic understanding.
- **Promote project-based learning:** Design assignments and projects that require students to apply both hard and soft skills to simulate real-world engineering challenges.

6.2 Faculty Training

- **Offer professional development:** Provide opportunities for faculty to update their knowledge and skills in emerging areas such as IoT, data analytics, and advanced control systems.
- **Encourage industry collaboration:** Facilitate partnerships between faculty and industry professionals to ensure that course content remains relevant and aligned with current industry practices.

- **Enhance teaching methodologies:** Train faculty in effective teaching methodologies that promote interdisciplinary learning, teamwork, and problem-solving.

6.3 Accreditation

- **Update accreditation standards:** Revise accreditation standards to include competencies related to emerging technologies, interdisciplinary skills, and professional skills such as communication and leadership.
- **Assess learning outcomes:** Implement assessment methods that evaluate students' ability to apply both hard and soft skills in real-world contexts, ensuring that graduates are well-prepared for the demands of modern engineering roles.
- **Promote continuous improvement:** Encourage programs to continuously evaluate and update their curriculum based on feedback from industry partners and alumni to ensure ongoing relevance and effectiveness.

Conclusion

This methodological framework integrates robust keyword extraction, advanced n-gram analysis, and rigorous qualitative validation. The resulting annotated dataset offers significant insights into workforce demands and skill trends, serving as a foundation for high-impact research and practical applications.

References

[1] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993-1022, 2003.

[2] L. K. H. Yanaze and R. de Deus Lopes, "Transversal competencies of electrical and computing engineers considering market demand," in *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*, Oct. 2014, pp. 1-4, IEEE.

[3] S. Lappin and S. M. Shieber, "Machine learning theory and practice as a source of insight into universal grammar," *Journal of Linguistics*, vol. 43, no. 2, pp. 393-427, 2007.

[4] V. Braun and V. Clarke, "Using Thematic Analysis in Psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77-101, 2006.

[5] M. Zhang, K. N. Jensen, R. van der Goot, and B. Plank, "Skill extraction from job postings using weak supervision," *arXiv preprint arXiv:2209.08071*, 2022.

M. Zhang, K. N. Jensen, S. D. Sonniks, and B. Plank, "SkillSpan: Hard and soft skill extraction from English job postings," *arXiv preprint arXiv:2204.12811*, 2022.

(Two references in list, only one is correct - as mentioned only once in text)

[6] T. A. R. B. T. Abdullah and N. I. B. Hamzah, "Development of a facility for an advance competency in Electrical Engineering," in *2009 3rd International Conference on Energy and Environment (ICEE)*, Dec. 2009, pp. 469-473.

[7] A. McGettrick, R. May, and R. Ward, "Principles underlying the development of competencies for engineers," *Computing & Control Engineering Journal*, vol. 11, no. 1, pp. 40-43, 2000.

[8] C. J. Leslie, "Engineering Competency Model," 2016

(Check this citation once, could not find complete info/details on google scholar)

[9] R. Szydlo, S. Wisniewska, M. Tyranska, A. Dolot, U. Bukowska, and M. Koczynski, "Competencies Demand From Young Engineers on the Energy Market in Poland," in *International Scientific and Practical Conference "Young Engineers of the Fuel and Energy Complex: Developing the Energy Agenda of the Future" (EAF 2021)*, Mar. 2022, pp. 207-216, Atlantis Press.

[10] D. Permatasari and D. Mardapi, "Competency Profile of Electrical Engineering Graduates Relevant to the Industry," in *Proceedings of the 2nd International Conference of Science and Technology for the Internet of Things, ICSTI 2019*, 3 Sep. 2019, Yogyakarta, Indonesia, Mar. 2020.

[11] C. Y. van Staden and K. S. Garner, "Perceived Practical Competency amongst Electrical and Electronic Engineering Students," in *2023 32nd Annual Conference of the European Association for Education in Electrical and Information Engineering (EAEEIE)*, Jun. 2023, pp. 1-6, IEEE.

[12] O. Ozyurt, F. Gurcan, G. G. M. Dalveren, and M. Derawi, "Career in cloud computing: exploratory analysis of in-demand competency areas and skill sets," *Applied Sciences*, vol. 12, no. 19, p. 9787, 2022.

[13] E. Moustroufas, I. Stamelos, and L. Angelis, "Competency profiling for software engineers: literature review and a new model," in *Proceedings of the 19th Panhellenic Conference on Informatics*, Oct. 2015, pp. 235-240.

[14] R. E. Boyatzis, *Transforming Qualitative Information: Thematic Analysis and Code Development*. Sage Publications, 1998.

[15] M. D. Mumford, M. A. Marks, M. S. Connelly, S. J. Zaccaro, and R. Reiter-Palmon, "Development of leadership skills: Experience and timing," *The Leadership Quarterly*, vol. 11, no. 1, pp. 87-114, 2000.

[16] L. B. Sørensen and G. Coates, "Assessing Technical and Professional Skills of Engineering Students: A Case Study," *European Journal of Engineering Education*, vol. 45, no. 1, pp. 79-94, 2020.

[17] M. M. Robles, "Executive perceptions of the top 10 soft skills needed in today's workplace," *Business Communication Quarterly*, vol. 75, no. 4, pp. 453-465, 2012.

[18] E. A. Fleishman, D. P. Costanza, and J. Marshall-Mies, "Abilities," in *The Changing Nature of Performance: Implications for Staffing, Motivation, and Development*, D. R. Ilgen and E. D. Pulakos, Eds. San Francisco, CA: Jossey-Bass, 1999, pp. 23–46

[19] D. N. Perkins, E. Jay, and S. Tishman, "Beyond abilities: A dispositional theory of thinking," *Merrill-Palmer Quarterly*, vol. 39, no. 1, pp. 1-21, 1993.

[20] Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. Sage Publications.

[21] Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. International Journal of Qualitative Methods, 16(1), 1-13.

Appendix A1

Removing Stopwords from and Defining Mappings for Standardizing Job Titles

```
[ ] # ======================================== RECONSTRUCT JOB TITLES ================================

    # Defining mappings for standardizing job titles
    title_mapping = {
        'Electrical Apprentice': 'Electrical Apprentice',
        'Electrical': 'Electrical Engineer',
        'Engineer I': 'Electrical Engineer',
        'Controls Engineer - job posts': 'Controls Engineer',
        'Controls Engineer-': 'Controls Engineer',
        'Director of Electrical Engineering': 'Director of Electrical Engineering',
        'Electrical Maintenance': 'Electrical Maintenance Engineer',
        'Vice President of Electrical - Solar': 'Vice President of Electrical Engineering',
        'Maintenance, Electromechanical': 'Electromechanical Maintenance Engineer',
        'Distribution Designer I or II': 'Electrical Distribution Designer',
        'Component Engineer - Hardware Engineering': 'Hardware - Component Engineer',
        'Engineering Technician I-III Traffic Safety Officer': 'Engineering Technician',
        'Junior - Team Lead Electrical Engineering': 'Junior Team Lead Electrical Engineer',
        'Controls Engineering Co-Op': 'Controls Engineer',
        'Electrician Journeyman': 'Electrician',
        'Design, Verification, and Validation Engineers': 'Design, Verification, and Validation Engineer',
        'Electrical Engineer 1': 'Electrical Engineer I',
        'Entry Level Electrical Engineering Graduate': 'Electrical Engineer',
        'ASIC Engineer, Design Verification': 'ASIC Engineer',
        'Entry Level Electrical Controls Engineer': 'Electrical Controls Engineer',
        'Water Plant Automation Controls Engineer': 'Automation Controls Engineer',
        'Entry Level Electrical Engineering Graduate': 'Electrical Engineer',
```

Appendix A2

```
# =============================================================== REMOVE STOPWORDS ===============================================================
# Define custom stopwords
custom_stopwords = ['job description', '8362', '8362', '4727', '37101', '37449', '19701', '160130', '002064', 'az001', 'duluth', 'dojo',
                    'edwardsville', 'ks', 'kansas', 'osc', 'adas', '3000', '21223', '1185', 'value', 'tennessee', 'ddr', 'cuas', 'core', 'hi', 'ent',
                    'earlymidcareer', 'driver', 'diffusion', 'ref1562c', 'radiant', 'q3', 'field002064', 'id', 'advisor', 'able', '8', '3pm', '2l', '1000',
                    'l', 'kit', 'iiiiiiivvvi', 'fsm', 'december', 'chicago', '12', '10k', 'ot', 'ncg', 'multiyear', 'ms', 'lithium', 'per', 'pe', 'ny', 'nm',
                    'mec', 'londonderry', 'hour', 'hospital', 'hill', 'hampshir', 'sta', 'smyrna', 'ii', 'iii', '-', 'job', 'post', '2nd', 'usa', 'staff', 'shift',
                    'sr', 'bess', 'multiple', 'updated', 'positions', 'rep', 'director', 'entry', 'level', 'mid', '0801', '14', 'gs', '10118','direct', 'hire',
                    'city', 'senior', 'physical', 'regional', 'hud', 'site', 'osp', 'remote', 'mep', 'multi', 'sw', 'fort', 'fork', 'greely', 'certification',
                    '110373', 'american', 'george', 'kiewit', 'line', 'bonus', 'regional', 'superintendent', 'building', 'gc', 'junior', 'mechanical', 'fire',
                    'years', 'tx', 'orem', 'principal', 'robot', '1318', 'sc', 'alaska', 'wilson', 'center', 'cost', 'supv', 'ellisville', 'drivability',
                    'conversions', 'bmet1', 'bmetl', '26', 'ux', 'physicist', 'co', 'op', 'fall', 'new', '1st', '4wd', 'engrg', 'commodity', 'jd2', 'pwb', 'ims',
                    'authority', 'environmental', 'mn', 'college', 'grad', 'warfare', 'iv', 'hybrid', 'career', 'early', '2024', 'emc', 'spring', 'tn', 'day',
                    'overall', 'third', 'se', 'ford', '90', '32', '30', '000', 'ote', 'cng', 'ct2', 'grad', 'improvement', 'continuous', 'fielder', 'harness',
                    'program', 'northwest', 'vp', 'civil', 'region', 'full', 'time', 'san', 'antonio', 'texas', 'vt', 'sign', '4831', 'chemical', 'hw', 'graduate',
                    'pathways', 'recent', 'survivability', 'md', 'electro', 'locations', '23', 'plu', 'vacancies', 'miamisburg', 'smr', 'rus', 'staking', 'beverage',
                    'food', 'january', 'semester', 'month', 'mill', 'august', 'machinining', 'Entry', 'Level', '(Multiple', 'Openings)', '(Level', '1', '2)', 'Shift']

# Extend the custom stopwords set
additional_stopwords = ['8362', '8362', '4727', '37101', '37449', '19701', '160130', '002064', 'az001', 'duluth', 'dojo',
                        'edwardsville', 'ks', 'kansas', 'osc', 'adas', '3000', '21223', '1185', 'value', 'tennessee', 'ddr', 'cuas', 'core', 'hi', 'ent',
                        'earlymidcareer', 'driver', 'diffusion', 'ref1562c', 'radiant', 'q3', 'field002064', 'id', 'advisor', 'able', '8', '3pm', '2l', '1000',
                        'l', 'kit', 'iiiiiiivvvi', 'fsm', 'december', 'chicago', '12', '10k', 'ot', 'ncg', 'multiyear', 'ms', 'lithium', 'per', 'pe', 'ny', 'nm',
                        'mec', 'londonderry', 'hour', 'hospital', 'hill', 'hampshir', 'sta', 'smyrna']

# concatenate the two custom stopwords
concat_stopwords = custom_stopwords + additional_stopwords

# # Get NLTK's English stopwords and combine the default with the custom stopwords
# combined_stopwords = set(ENGLISH_STOP_WORDS).union(custom_stopwords_list)
combined_stopwords = set(stopwords.words('english')) #.union(custom_stopwords) # stopwords.words('english')
# update the custom stopwords
combined_stopwords.update(concat_stopwords)
combined_stopwords = list(combined_stopwords)
```

Appendix B

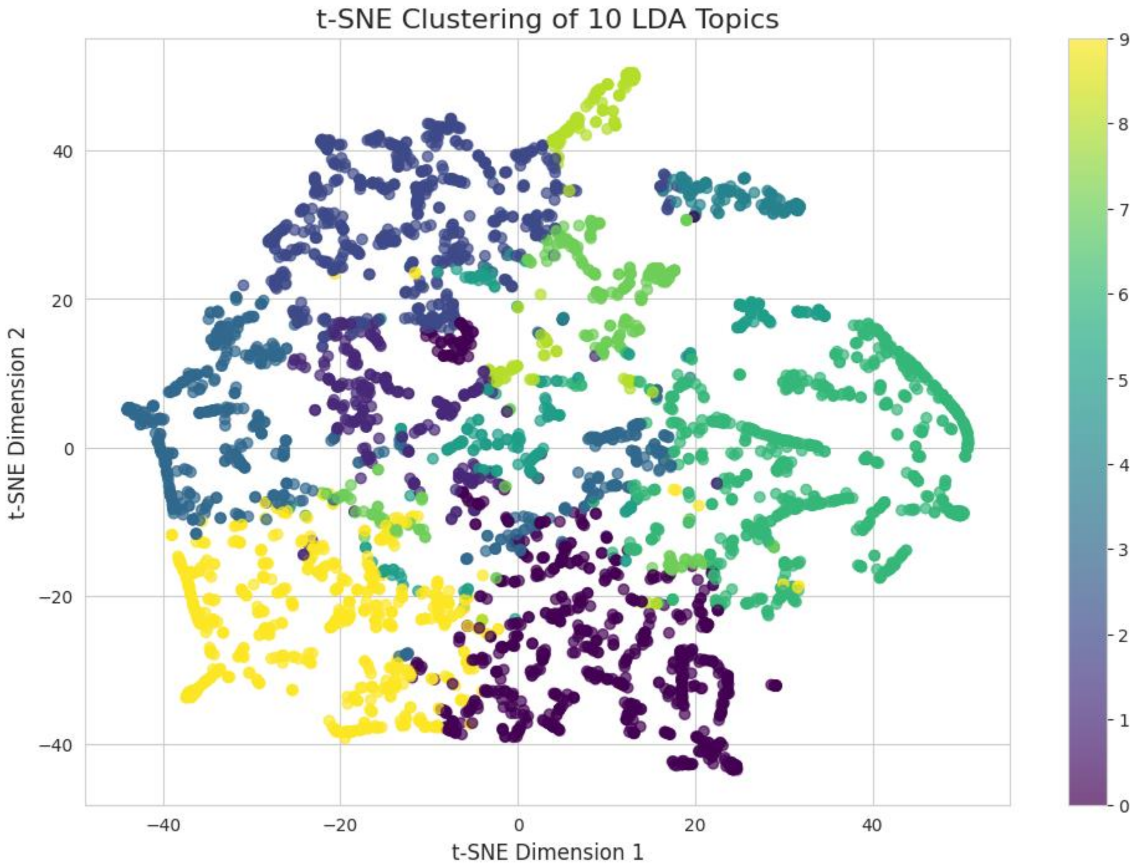

Figure 8. Visualizing Thematic Clustering of LDA Topics via t-SNE Clustering

Appendix C

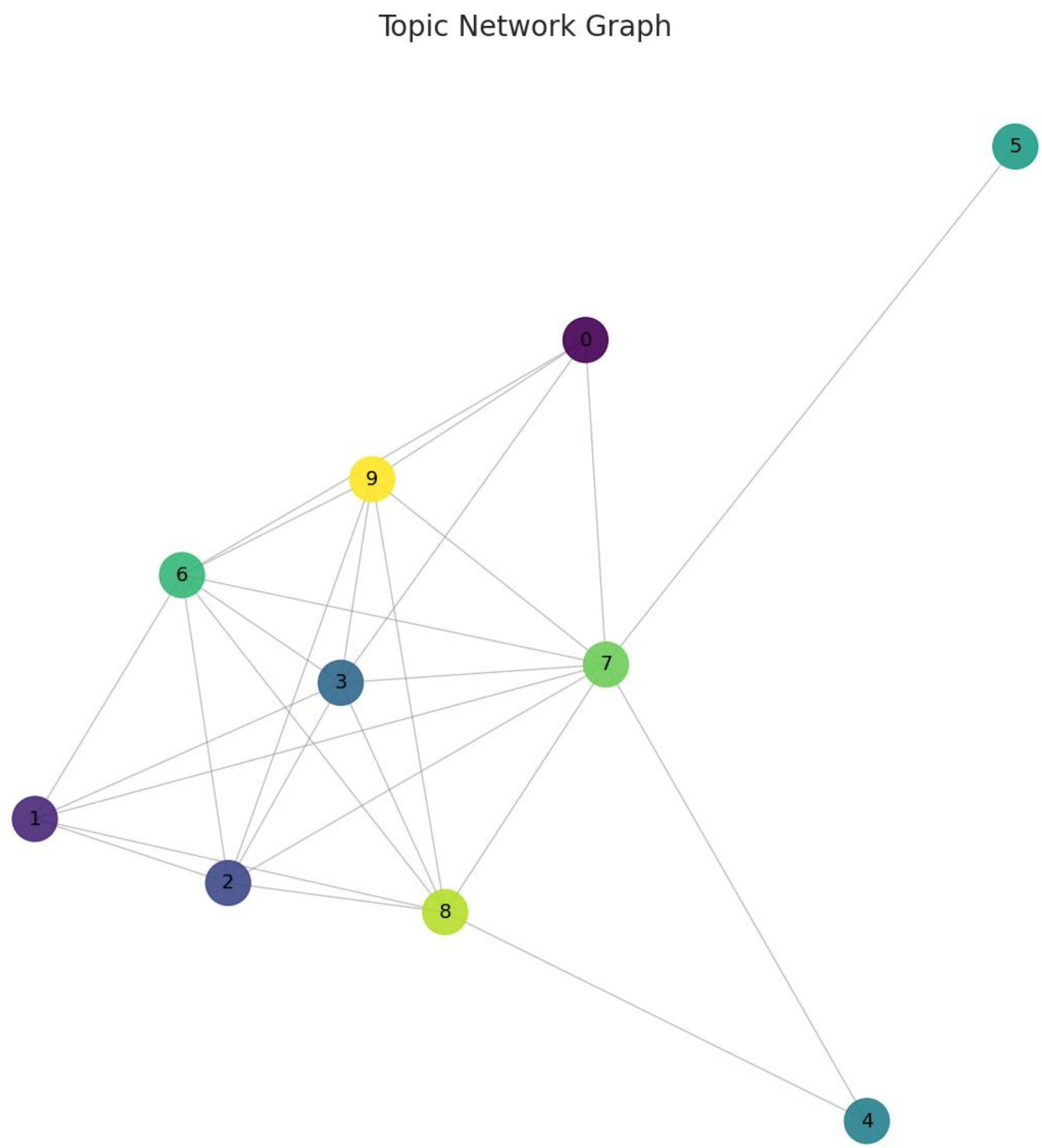

Figure 9. Visualizing Thematic Interconnections of LDA Topics via a Topic Network Graph