Print ISSN: 1738-3110 / Online ISSN 2093-7717
JDS website: http://accesson.kr/jds
http://doi.org/10.15722/jds.24.03.202603.1

Data-Driven Labor Market Analytics for Diagnosing Workforce Constraints in Retail and Distribution Systems

Leyla GAMIDULLAEVA[1], Mikhail DEEV[2], Anel KIREYEVA[3], Alexey FINOGEEV[4], Svetlana KOZHIROVA[5]

Received: January 14, 2026. Revised: February 12, 2026. Accepted: March 05, 2026.

Abstract

Purpose: This study examines competence gaps as a structural constraint hindering retail and distribution systems' adaptation to digitalization and omnichannel models. Departing from a pedagogical perspective, it treats workforce competencies as an operational resource of these systems. Consequently, educational programs are analyzed not as a primary object of study, but as an institutional benchmark and proxy for interpreting competence demand within retail and distribution systems. **Research design, data and methodology:** The analysis uses vacancy data (approximately 1.2 thousand postings from the Penza regional market, 2024) for retail, logistics, and SCM occupations. Applying semantic text analysis and machine learning (NLP, clustering), skills were aggregated into competence profiles to identify misalignments between employer demand and institutional benchmarks from university programs. **Results:** Results show that modern systems rely on hybrid competence configurations combining digital, analytical, managerial, and customer-oriented components. These configurations are linked to inventory coordination, logistics responsiveness, and omnichannel service quality but remain poorly reflected in institutional benchmarks. **Conclusions:** The study demonstrates that vacancy-based labor market analytics can serve as a distribution-oriented diagnostic tool for identifying workforce-related coordination constraints. The proposed framework provides a data-driven basis for strengthening adaptive capacity in retail and distribution systems.

Keywords: Distribution, Distribution System, Labor Market Analytics, Retail and Distribution Management, Supply Chain Workforce, Educational Program Alignment

JEL Classification Code: J24; M53; L81

1. Introduction

Modern retail and distribution systems operate in an environment characterized by rapid digital transformation, increasing supply chain complexity, and the expansion of omnichannel business models. Under these conditions, workforce competencies constitute a critical operational resource of distribution systems, alongside physical

1 First Author. Doctor of Economic Sciences, Professor, Penza State University, Penza, Russia. Email: gamidullaeva@gmail.com
2 Second Author. PhD candidate, Penza State University, Penza, Russia. Email: miqz@yandex.ru
3 Corresponding Author. PhD, Professor, Department of Scientific and Research Activities, University of International Business, Almaty, Kazakhstan. Email: kireyeva.anel@ieconom.kz
4 Fourth Author. Doctor of Technical Sciences, Professor, CAD Department of Penza State University, Penza, Russia. Email: alexeyfinogeev@gmail.com
5 Fifth Author. PhD, Professor, International Science Complex Astana, Astana, Kazakhstan. Email: svetlana.kozhirova@spr.kz

ⓒ Copyright: The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://Creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted noncommercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

infrastructure and logistics technologies. Today, effective distribution networks require not only infrastructure and technology, but also a qualified workforce capable of managing data-driven operations, optimizing inventory, and interacting with customers across integrated channels. Recent studies emphasize that digitalization and analytics are becoming critical factors shaping the performance of retail and distribution systems (Tzimas et al., 2024).

In recent years, retail chains, wholesale intermediaries, and logistics service providers have increasingly struggled to recruit specialists whose qualifications align with the tasks of modern distribution systems. Positions related to distribution are no longer limited to narrowly defined functional roles; instead, they presuppose the formation of hybrid competence profiles. Such profiles combine skills in data analysis and digital technologies with managerial decision-making and practical experience in working with customers and distribution partners. When such hybrid competence configurations are insufficiently available, they emerge as structural bottlenecks constraining distribution system performance. Insufficient alignment between workforce competencies and evolving distribution tasks directly constrains the operational performance of retail and distribution systems. In practice, this is reflected in weaker inventory coordination, lower logistics responsiveness, and declining service quality across integrated channels. This problem has been widely discussed in research on labor market monitoring and education–employment alignment, particularly in service-oriented spheres (OECD, 2021; European Commission, 2021; Beresewicz et al., 2025). However, existing studies rarely treat workforce competence mismatches explicitly as a distribution-system problem or analyze them through the lens of operational performance and coordination.

The problem of aligning educational programs with labor market demand has become increasingly acute amid digital transformation. From a distribution-system perspective, this issue should be interpreted not as a pedagogical challenge, but as an institutional lag in the reproduction of competence structures demanded by retail and distribution systems. Traditional mechanisms for competence structure adjustment, such as professional standards, state educational frameworks, and expert-based revisions, are primarily unable to respond adequately to the pace and nature of ongoing structural changes. As noted in earlier studies, these instruments tend to overlook emerging competence requirements and hybrid skill configurations characteristic of rapidly transforming sectors (Mocanu et al., 2014). In retail and distribution systems, such institutional inertia manifests itself in reduced workforce adaptability and limits the capacity of distribution networks to respond to technological and organizational change.

Existing labor market analysis tools also have important limitations. Aggregated statistics and periodic surveys provide only a general view of workforce demand and rarely reflect the specific competence profiles employers require. In contrast, vacancy texts provide detailed, dynamic information on the skills, responsibilities, and technologies in demand. Recent research demonstrates that automated analysis of job postings using machine learning and natural language processing enables more accurate and timely identification of labor market requirements (Botov et al., 2019; Chen et al., 2019). Nevertheless, these analytical advances are still rarely embedded in distribution-oriented diagnostic frameworks that interpret competence demand as a constraint on system performance rather than as an isolated labor market signal.

The persistent separation between labor market analytics and curriculum development represents one of the principal limitations of existing research approaches. From a distribution-science standpoint, this separation limits the ability to diagnose competence mismatches as sources of coordination inefficiency within distribution systems. Although a considerable body of studies focuses on vacancy-based skill analysis and labor demand forecasting, substantially less attention has been given to methodological frameworks that translate these analytical insights into system-level interpretations relevant to retail and distribution management. In retail and distribution management, the limits of existing competence frameworks are often revealed at the operational level. The growing use of omnichannel sales models, digital logistics solutions, and data-supported managerial practices changes how employees combine tasks and skills in their daily work. However, these changes are usually included in standardized occupational descriptions only after a significant delay.

In retail and distribution systems, this lag is particularly consequential given the high degree of operational interdependence and coordination required across logistics, inventory management, and customer-facing processes. The widespread adoption of logistics analytics tools, automated warehouse solutions, and digitally integrated customer service functions leads to frequent changes in job content and competence requirements. Under such conditions, institutional representations of competence demand tend to adjust with delay, particularly in the absence of stable analytical feedback mechanisms linking distribution-system requirements with institutional benchmarks.

In this context, the present study addresses the need for a dynamic, data-driven approach to diagnosing workforce constraints in retail and distribution. The paper proposes and tests a methodology for analyzing labor market competencies using automated collection of vacancy data and machine learning techniques. The analytical focus of the

study is the retail and distribution labor market; educational programs are used solely as institutional benchmarks reflecting how competence demand generated by distribution systems is formalized and reproduced.

Accordingly, this research does not examine education as an autonomous domain and does not aim to contribute to pedagogical theory, curriculum design, or educational policy. Instead, educational program descriptions are employed as comparative artifacts to assess the degree to which institutional competence structures correspond to the evolving functional requirements of distribution systems.

From a methodological perspective, the study contributes an integrated analytical framework that links vacancy-based labor market data with the benchmarking of institutional competence structures. From a distribution science perspective, the contribution lies in conceptualizing workforce competence as a system resource, treating competence misalignment as a form of distribution inefficiency, and demonstrating how vacancy-based analytics can support adaptive coordination in retail and distribution systems. The practical significance of the results lies in their applicability for retail and distribution organizations and public authorities concerned with workforce constraints under digital logistics and omnichannel operating models, while educational programs appear as a secondary, derivative domain of interpretation.

This study investigates workforce competence mismatches as structural bottlenecks that constrain the functioning and adaptation of retail and distribution systems undergoing digital and omnichannel transformation. We frame these mismatches not as isolated educational shortfalls, but as systemic constraints hampering inventory coordination, logistics responsiveness, and service performance across distribution networks. Consequently, we employ labor market analytics as a diagnostic instrument for these systems, using educational programs solely as institutional benchmarks to evaluate how the competence demand generated by distribution systems is formalized within existing frameworks.

2. Literature Review

The alignment of workforce competencies with labor market requirements has been examined across several strands of research, particularly in relation to digital transformation and structural change in retail and distribution systems (Assante et al., 2019; Trenerry et al., 2021). Modern distribution networks increasingly depend not only on physical infrastructure but also on the availability of qualified personnel capable of operating complex logistics systems, managing omnichannel retail formats, and using data-driven decision-making tools. From the perspective of distribution science, workforce competencies are increasingly conceptualized as a critical operational resource that directly influences system performance metrics, such as inventory turnover, logistics responsiveness, and supply chain resilience (Waller & Fawcett, 2013; Singh, 2025). Consequently, competence mismatches are not merely a labor market issue but a source of systemic inefficiency that can constrain coordination and adaptive capacity within distribution networks. This problem has been widely discussed in research on labour market monitoring and education–employment alignment, including work on vacancy-based measurement and the institutionalization of workforce analytics (Beresewicz et al., 2025; Coolen et al., 2023). As a result, the gap between competence demand and supply has become a critical constraint on the efficiency and resilience of retail and distribution sectors.

2.1. Labor Market Analytics and Vacancy-Based Skill Analysis

An examination of the research literature shows that, despite the rapid expansion of labor market analytics and the growing use of advanced analytical instruments – including machine learning-based approaches to skill forecasting and computational analysis of job postings, as well as studies on education–employment alignment – a number of substantive issues remain insufficiently explored (e.g., Bakhshi et al., 2017; OECD, 2017; Khaouja et al., 2021; Senger et al., 2024). In particular, existing works rarely integrate automated vacancy-based skill analysis with a consistent and structured evaluation of educational programs, especially in sector-specific contexts such as retail and distribution (Kureková et al., 2015; Deming & Noray, 2020).

At the same time, competence requirements associated with omnichannel retail formats, digitally enabled logistics, and increasingly complex distribution systems are highly dynamic, while prevailing analytical frameworks still reflect these changes only fragmentarily. As a result, insights derived from vacancy analytics are rarely translated into real-time diagnostic indicators of workforce bottlenecks that constrain inventory coordination, logistics responsiveness, and overall system adaptability (Tavakoli et al., 2022). Similar limitations have been identified in studies, which emphasize that even when vacancy-based analytics successfully capture emerging skill requirements, their translation into systematic tools for curriculum benchmarking and workforce planning at the sectoral level remains methodologically underdeveloped (McGuinness et al., 2018).

Against this background, the present study advances an integrated methodological approach that combines automated labor market data collection, machine learning–

based analysis of competence demand, and comparative evaluation of educational programs. With a specific focus on retail and distribution management, the research contributes to distribution science by clarifying the linkage between labor market analytics and the formation of competence profiles that support the effective functioning and further evolution of modern distribution systems.

2.2. Machine Learning and NLP in Competence Modeling

Methods in machine learning and natural language processing have become integral to analytical approaches for interpreting unstructured vacancy texts and translating them into structured representations of skill and competence demand (Khaouja et al., 2021; Senger et al., 2024). Early research highlighted the potential of vector-based semantic models for identifying and comparing skills reflected in labor market data. Foundational contributions to this research stream include distributional word representation models, which enabled semantic comparison beyond explicit keyword overlap (Mikolov et al., 2013; Le & Mikolov, 2014). Building on these ideas, Mocanu et al. (2014) demonstrated that NLP-based analysis of vacancy texts can support the identification of emerging competence requirements and inform adjustments in educational curricula.

Subsequent studies further operationalized these approaches by applying ontology-based and text-mining methods to large-scale vacancy datasets, improving the granularity and consistency of skill extraction and enabling systematic mapping between vacancy texts and formal competence frameworks (Sibarani et al., 2017; Miranda et al., 2017; Dadzie et al., 2018; Spada et al., 2024). Alongside these methods, embedding-based techniques have been employed to capture latent semantic relationships among skills and competence elements, facilitating the identification of stable links between skill descriptions in job advertisements and competence structures embedded in competence frameworks and related standards (Mikolov et al., 2013; Le & Mikolov, 2014; Senger et al., 2024). A key advantage of such representations is their ability to move beyond surface-level keyword matching and to reveal latent connections between different groups of competencies.

A frequently cited illustration of this approach is the study by Botov et al. (2019), which develops a procedure to compare professional standards with vacancy texts using distributed semantic representations and neural network models. The authors show that vector-based methods provide a more reliable basis for detecting discrepancies between educational content and labor market requirements than conventional keyword-based techniques. This methodological approach is particularly relevant for retail and distribution management, where competence profiles are rarely limited to single skill categories. In practice, job roles in these sectors typically combine technical, analytical, and customer-oriented elements, which are difficult to describe using rigid occupational classifications. Subsequent studies have further developed this line of research by applying machine learning and deep learning techniques to large and heterogeneous datasets of job advertisements. These studies show that neural network–based models are better suited for extracting, clustering, and classifying skills than rule-based or keyword-driven methods, especially when skill requirements are expressed in composite or context-dependent forms (Boselli et al., 2018; Khaouja et al., 2021). This advantage is particularly important for retail and distribution systems, where job design often integrates analytical tasks, managerial responsibilities, and service-related functions within the same role, making surface-level text matching insufficient for capturing actual competence requirements.

2.3. Competence Demand as a Constraint on Distribution System Performance

Research on labor market dynamics increasingly examines how structural mismatches between competence supply and demand constrain organizational and systemic performance, particularly in contexts of rapid technological and organizational change (OECD, 2017; McGuinness et al., 2018). A significant portion of this literature investigates the manifestation of these mismatches through the lag of institutional training frameworks behind the evolving competence requirements of industries. In particular, Tavakoli et al. (2022) analyzed personalized learning systems that adapt educational trajectories based on labor market signals, highlighting the importance of curricular flexibility amid accelerated economic change. Related perspectives are developed within the literature on competence-based education and modular program design, where consideration of employer demand is seen as key to maintaining the relevance of learning outcomes.

At the same time, numerous studies indicate that traditional educational standards continue to lag behind the rapidly changing labor market requirements, especially in sectors undergoing intensive digital transformation. Vacancy-based evidence and skill analytics increasingly document gaps in applied competencies – such as data analysis, automation, project management, and service design – that are critical for digitally enabled retail and distribution environments (Pejić-Bach et al., 2021). These deficiencies are particularly consequential for retail chains, distribution centers, and logistics operators, where analytical capabilities and digital decision-support skills directly influence the efficiency and reliability of

distribution processes. In distribution-intensive contexts, data-driven decision-making and analytical competencies are frequently identified as important determinants of operational performance (Waller & Fawcett, 2013), while customer-oriented and service-related skills remain essential components of hybrid competence profiles (Grönroos & Voima, 2013). This systemic competence mismatch, evidenced by the lag of institutional frameworks, contributes to structural bottlenecks in employment and directly constrains the operational performance, coordination capacity, and resilience of distribution systems.

Despite the acknowledged relevance of these challenges, much of the existing research remains either overly general or limited to individual industries, without sufficient consideration of the specific functional requirements of distribution networks (Senger et al., 2024). In many cases, labor market dynamics are examined retrospectively, and analytical approaches do not provide instruments for continuous monitoring or for the timely adjustment of educational content. Moreover, the integration of vacancy-based machine learning analysis with systematic benchmarking of educational programs remains insufficiently developed, particularly in sector-specific domains such as retail and distribution management.

2.4. Research Gap and Contribution of the Study

The analysis of existing research shows that, despite significant advances in labor market analytics and machine learning-based competence modeling, a critical diagnostic gap persists in the field of distribution science. In particular, the literature offers limited methodological frameworks that treat real-time vacancy data as a systemic indicator for diagnosing competence-related bottlenecks that hinder the operational efficiency, digital adaptation, and resilience of retail and distribution networks.

While recent publications increasingly discuss digital transformation and the use of Big Data analytics in logistics and retail companies (Kim et al., 2024; Kitcharoen, 2023), the dynamic nature of competence requirements generated by omnichannel retail models, digital logistics platforms, and growing supply chain complexity is still insufficiently reflected in prevailing analytical frameworks.

Moreover, the translation of labor market analytics into diagnostic tools for identifying systemic bottlenecks in distribution networks remains fragmented and methodologically underdeveloped. Although distribution research offers extensive evidence on data-driven tools for operational performance management, these advances are rarely transformed into structured approaches for diagnosing workforce competence constraints that affect logistics coordination and omnichannel integration (Kitcharoen, 2023; Jing & Fan, 2024; Jiang et al., 2025; Singh, 2025).

Against this background, the present study aims to address this diagnostic gap by developing and testing a methodological framework that utilizes automated labor market data collection and machine–learning–based competence modeling to identify and quantify structural competence misalignments. By focusing on retail and distribution systems, the study contributes to distribution science by conceptualizing workforce competence as a key system resource and by providing a data-driven diagnostic tool to pinpoint competence gaps that act as bottlenecks for inventory coordination, logistics responsiveness, and omnichannel service integration. Educational programs are utilized not as a primary object of reform, but as a stable institutional benchmark against which the evolving demands of the distribution sector can be measured.

3. Research Methods

3.1. Functions of the Competence Selection, Comparison, and Analysis System

The competence analysis system serves as a diagnostic tool. It identifies and quantifies structural competence mismatches in retail and distribution systems. Its primary function is to benchmark the competence demand articulated in job vacancies - a real-time signal generated by the evolving needs of distribution networks - against institutionalized competence structures, such as those embedded in institutional competence representations. This benchmarking does not treat education as a primary domain of intervention but uses it as a stable reference framework to assess how well institutionalized competence profiles reflect the current operational requirements of distribution systems.

The term “skill” in this study refers to individual abilities or keywords extracted from job vacancy texts. By contrast, “competence” refers to an integrated functional capability that emerges from the aggregation and clustering of multiple skills and is applied for benchmarking purposes.

From a practical perspective, the proposed system is designed to support several closely related analytical functions. First, it enables the identification of current and emerging employer requirements based on labor market data. Second, it provides an analytical basis for examining competence demand in management, logistics, and distribution. Finally, the system structures and standardizes heterogeneous competence requirements, creating a foundation for comparative analysis of institutional competence representations that train specialists for retail chains, wholesale distribution, and logistics companies.

This reasoning aligns with studies that view universities as active participants in cross-industry knowledge ecosystems. In this perspective, digital analytical tools support ongoing interaction between labor market signals and the design of educational programs (Tolstykh et al., 2021).

3.2. The Method of Searching, Processing, and Consolidating Competence Requirements

The proposed methodology is based on a step-by-step transformation of labor market information into competence profiles suitable for analytical interpretation. These profiles are subsequently used for institutional benchmarking, with institutional profiles serving as comparative reference structures. The overall logic of the process is illustrated in Figure 1 and reflects the transition from unstructured job vacancy data to structured competence representations that enable systematic comparison.

At the initial stage, labor market data are collected from open-access sources containing job vacancy descriptions related to retail, logistics, and distribution activities. The empirical analysis draws on approximately 1.2 thousand online job vacancy postings from the Penza regional labor market collected in 2024, together with educational programs implemented by universities in the Penza region. We collected vacancy data using the Rosnavyk labor market analytics service. This service integrates data from several major online recruitment platforms, such as HeadHunter and SuperJob. The analysis covers a broad range of job advertisements, including logistics managers, supply chain specialists, warehouse supervisors, purchasing managers, and inventory planners.

We filtered the initial dataset to remove duplicate postings, incomplete vacancy descriptions, and announcements unrelated to distribution-related occupations. Only vacancies containing explicit descriptions of job tasks and competence requirements were retained for further analysis. This procedure was used to improve the consistency of the vacancy dataset and to ensure that the analyzed postings were relevant to retail and distribution activities.

The resulting dataset lets us examine employer demand in job advertisements. We can observe how competence requirements vary across different job types. Vacancy texts also capture changes in job content related to the adoption of digital tools, automation, and omnichannel business practices within retail and distribution systems. These factors are reflected in the way tasks are described and in the combination of skills and competencies required by employers.

Before analysis, the collected vacancy data were consolidated and subjected to basic preprocessing. This included merging data from different sources, removing duplicate postings, and normalizing textual descriptions. As a result, a single cleaned corpus of vacancy texts was obtained and used for subsequent analytical procedures.

The next stage focuses on textual and semantic analysis of vacancy descriptions.

At this stage, unstructured textual information is transformed into structured representations of competence requirements. Key terms and expressions describing knowledge, skills, and abilities are identified and formalized as vector representations. These representations enable the tracing of semantic relationships among competencies and the construction of competence profiles that reflect employer demand in the distribution sector. We incorporate institutional competence structures into the analytical framework at the benchmarking stage (Figure 1). They provide a standardized baseline. We use this baseline to measure and interpret the emergent competence demand from the distribution sector.

Vacancy texts were cleaned, deduplicated, and structured using automated linguistic and matching procedures to ensure consistency of employer information across multiple data sources. Technical details of data collection and preprocessing are provided in Appendix A.

Unlike vacancy data, we do not collect or preprocess educational programs. Instead, we directly transform their competence descriptions into vector representations. We use the same semantic space as for labor market data. This ensures methodological comparability between employer requirements and educational content. For text vectorization, the study employs a distributed semantic embedding approach based on the Word2Vec model (skip-gram architecture), trained on the combined corpus of vacancy texts and educational program descriptions. The Word2Vec model was selected for its ability to capture stable semantic relationships in medium-sized, domain-specific corpora while preserving the interpretability of competence clusters, which is critical for educational benchmarking tasks.

At the final stage, clustering and comparative analysis are performed to diagnose the degree of structural alignment (or misalignment) between competence demand in the distribution sector and its prevailing institutional representations. The number of clusters is defined a priori based on the structure of institutional profiles and fields of study, ensuring that clusters are interpretable in terms of existing specializations rather than purely data-driven partitioning. As a result, it becomes possible to distinguish between cases of full alignment, partial alignment, and substantial mismatches between labor market requirements and educational program content. The analysis produces structured competence profiles and alignment categories that serve to pinpoint specific workforce-related bottlenecks and coordination gaps within retail and distribution systems.

The practical implementation of such benchmarking procedures presupposes the availability of a digital educational environment capable of integrating analytical outputs into curriculum design and learning outcomes. Previous studies have shown that web-based learning environments and big data–driven educational platforms provide the necessary infrastructure for embedding labor market analytics into higher education systems (Mkrttchian et al., 2019; Deev et al., 2021; Mkrttchian et al., 2021). The corresponding benchmarking logic and clustering procedure are described in Section 3.3.

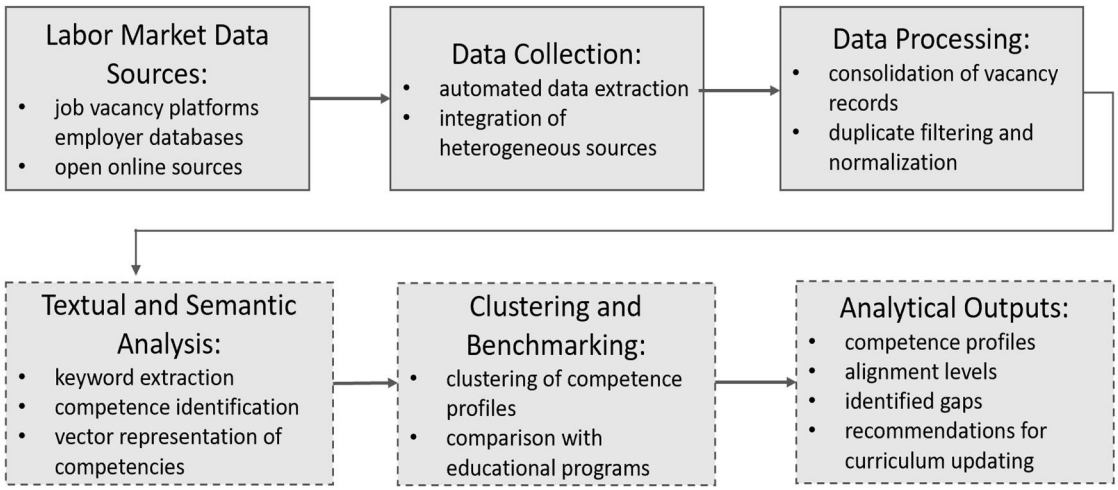

Figure 1: Research Framework for Labor Market Analytics and Educational Program Alignment

Figure 1 summarizes the analytical logic underlying the proposed diagnostic approach. It illustrates how vacancy-based competence extraction, clustering, and benchmarking are integrated to identify and map structural gaps between the competence demand generated by retail and distribution systems and their prevailing institutional representations.

The educational program dataset includes bachelor's and master's programs from universities in the Penza region. In total, curricula from several universities, including Penza State University and Penza State Technological University, were analyzed, comprising approximately 280 competency items extracted from officially documented institutional competence structures.

The analysis covers curricula valid in 2024 and examines declared learning outcomes, competence formulations, and skill requirements specified in educational standards and program descriptions. Individual competencies were decomposed into comparable elements to enable systematic benchmarking against vacancy-based competence profiles.

3.3. Fuzzy Clustering Method for Vectors of Competence Descriptions

Building on the competence profiles obtained through the procedures described in Section 3.2, the next stage of the analysis focuses on clustering and benchmarking vector representations of competencies in order to assess the alignment between educational programs and labor market demand in the distribution sector. The selection of clustering parameters was guided primarily by interpretability considerations rather than purely statistical optimization.

Benchmarking competence vectors provides a structured basis for identifying where the competence supply, as reflected in institutional profiles, deviates from the demand generated by the distribution sector. In the context of regional labor markets, such analysis enables diagnosing the severity and nature of competence misalignments, distinguishing between areas of strong congruence, partial mismatch, and significant structural gaps. The outcomes of benchmarking procedures provide a practical basis for revising educational content and for limiting the reproduction of competence profiles that are underrepresented in the actual distribution-related employment. Beyond a purely descriptive function, benchmarking allows the identification of areas where educational programs diverge from prevailing labor market requirements.

The use of clustering and benchmarking techniques is central to diagnosing competence-related constraints in retail and distribution systems, where competence requirements tend to be hybrid and context-dependent. In these settings, operational, analytical, managerial, and customer-oriented skills are integrated into cohesive job roles that directly influence system-level outcomes, such as

logistics flow and service integration. This interdependence reduces the analytical value of rigid classification schemes and increases the usefulness of vector-based benchmarking approaches, which allow competencies to be examined through their semantic relationships and patterns of co-occurrence instead of as isolated attributes.

The benchmarking procedure is designed to solve both direct and inverse analytical tasks. Competence vectors derived from vacancy descriptions are compared with vector representations of institutional competence structures, enabling the identification of the skills and areas of knowledge most in demand by employers. In the inverse setting, competence vectors of educational programs are compared with aggregated employer demand to assess the degree of program compliance with labor market requirements. Both tasks are implemented within a unified analytical framework based on automated clustering of vector representations.

Clustering is performed using predefined reference centers, depending on the direction of analysis. While fuzzy clustering approaches may be applied to competence analysis, in the present study the number of clusters is specified a priori, and reference competence vectors are fixed. This preserves interpretability and ensures a clear distinction between aligned and misaligned competence profiles, which is essential for educational program assessment. At the same time, this choice involves a methodological trade-off, as fixed reference centroids may limit the identification of emerging or unconventional hybrid competence profiles, which could be addressed in future research through adaptive or dynamic clustering strategies.

The analysis operates on several analytically distinct sets of competence vectors. These include vectors derived from formalized competence frameworks of regional universities, programs from neighboring regions, retrospective vacancy data, and current job vacancies. Retrospective and regional program vectors are used to establish reference cluster structures, while current vacancy vectors are benchmarked against these references. Vectors that do not align with regional reference clusters are subsequently compared with educational programs from neighboring regions, allowing the distinction between cases that require program updates and cases in which demand is satisfied by external educational supply.

From an analytical perspective, clustering of competence vectors serves not only as a technical tool for grouping similar profiles but also as a means of revealing latent structural relationships within the labor market. In the context of educational program assessment, such clustering makes it possible to move beyond binary judgments of relevance and instead identify transitional and borderline competence configurations. This, in turn, creates a basis for more flexible and differentiated decisions regarding curriculum updating and program positioning within the broader distribution labor market.

Following clustering and exponential normalization of semantic similarity between competence vectors, discrepancies between competencies embedded in educational programs and those required by employers are assessed. Three similarity intervals are defined to reflect high, moderate, and low levels of alignment between educational content and labor market demand. Programs demonstrating high similarity are interpreted as well aligned with current regional demand, while programs occupying an intermediate position are considered partially aligned and therefore require targeted updating.

Educational programs falling into the lowest similarity interval are characterize by a substantial mismatch with current regional labor market requirements. Such results indicate an increased risk of insufficient alignment between training outcomes and employer needs. At the same time, low alignment should not be interpreted unambiguously as evidence of irrelevance, as it may reflect temporary oversupply, deferred demand, or the spatial redistribution of labor demand.

Accordingly, programs assigned to this group are additionally analyzed in the context of neighboring regional labor markets.

For cases of partial alignment, the analysis identifies specific competence vector components where institutional profiles diverge from sectoral demand. These identified divergences highlight precise areas where the reproduction of competencies lags behind the operational needs of distribution systems. While these insights can be translated into recommendations for curricular adjustment, their primary value lies in providing a granular map of competence supply-demand frictions at the system level.

At the same time, targeted updating of educational programs based on identified competence gaps contributes not only to improving curriculum relevance but also to enhancing the overall quality of training and teaching practices. Such adjustments reduce the risk of forming competencies that remain weakly demanded in the labor market, a problem widely observed in institutional analyses of workforce reproduction.

Formal representation of the analytical procedure is provided in Appendix B, while technical details of model implementation are described in Appendix A. The overall architecture of the intelligent analysis system supporting competence clustering and benchmarking is illustrated in Figure 2. It illustrates the interactions among data collection, semantic processing, and analytical modules that translate heterogeneous labor market information into structured competence profiles.

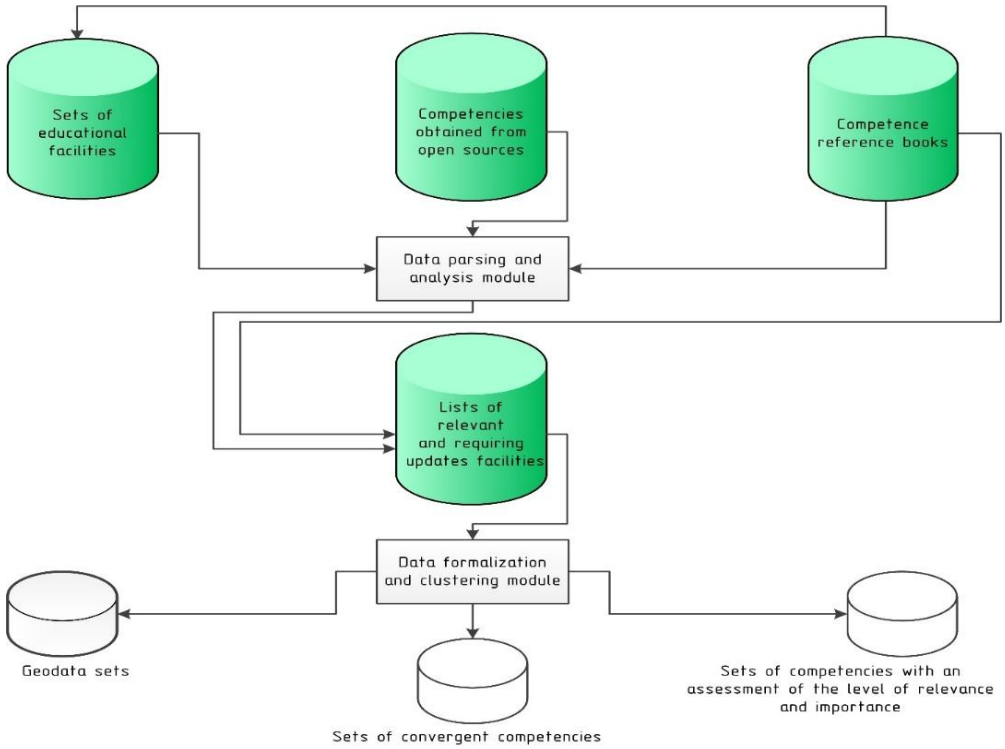

Figure 2: Modules of the System for Intelligent Analysis of Competence Vectors in Education Programs and Job Vacancies

In the empirical application, clustering was carried out with reference to the main fields of study covered by the analyzed institutionalized competence models. On this basis, four clusters were defined and subsequently used for benchmarking purposes. Benchmarking and clustering of competence vectors are used to compare education programs in management, logistics, and distribution with competence profiles extracted from vacancies in retail chains, distribution centers, and logistics companies. The model's output helps identify programs that are closely aligned with the needs of the distribution labor market, those that are partially aligned, and those that require substantial updating.

4. Results and Discussion

The proposed methodological approach was empirically assessed using labor market data from the Penza region. This empirical setting was selected because the region's employment structure is broadly representative of many Russian regions and exhibits an increasing influence of digitalization processes in trade, logistics, and distribution networks. Such a combination enables examination of workforce constraints relevant to the development of modern distribution systems under conditions of structural and technological transformation.

4.1. Application of the Competence Analytics Methodology Based on the Penza Region Labor Market

An examination of wage levels across major occupational groups indicates that the highest remuneration is observed in information technology positions (RUB 60,579), followed by engineering and management occupations (Figure 3). This wage configuration reflects a broader structural reorientation of labor demand toward competencies supporting the functioning and transformation of complex socioeconomic systems, rather than sector-specific characteristics alone. The figure shows differences in average wage levels across selected occupational groups, including digitally intensive and analytical positions relevant to retail and distribution systems. The application of the proposed methodology enabled the identification of the most in-demand competencies in the regional labor market, which are visualized as a word cloud (Figure 4).

He results reveal a clear predominance of analytical and digital competencies, including data analysis, programming, and information security. Within retail chains and distribution networks, these competencies form the functional foundation for developing electronic distribution channels, enabling real-time inventory management, and integrating logistics operations across multiple organizational levels. A comparative analysis of the most

and least demanded skills demonstrates a fundamental transformation in the structure of labor demand (Table 2).

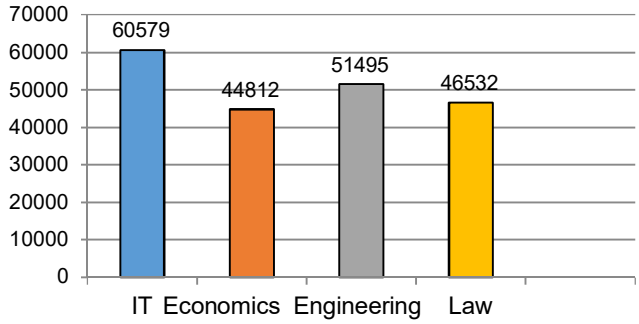

Figure 3: Average Salary of Specialists in Some Fields

Figure 4: Word Cloud of in-Demand Competencies

Table 2: Top 10 Most and Least Sought-After Skills

No.	Soft Skills	Technical Skills
Most sought-after skills		
1	Communication skills	Data analysis
2	Risk management	Programming (Python, Java, C++, SQL)
3	Analytical thinking	Cybersecurity
4	Creativity	Business process automation
5	Teamwork	Financial analysis
6	Customer orientation	Mobile app development
7	Adaptability	Project management
8	Self-organization	Cloud computing
9	HR management	Machine Learning & AI
10	Public speaking skills	Working with ERP systems (SAP, 1C: Enterprise, Oracle)
Least sought-after skills (baseline employer expectations)		
1	Politeness	Document management
2	Neatness	Office equipment operation
3	Friendliness	Accounting
4	Maintaining cleanliness and order	Interacting with government agencies
5	Desire to develop	Document verification
6	Literate speech	Workability
7	Responsibility	Working with archives
8	Punctuality	Processing sick leave forms
9	Stress resistance	Correspondence handling
10	Self-organization	General knowledge of legal acts

Note: Skills listed in the lower-ranked block reflect baseline emp loyer expectations that are usually taken for granted and therefo re mentioned less explicitly in job vacancy descriptions. Among the most frequently mentioned skills, technical skills related to d ata analysis, ERP systems, automation are particularly relevant for distribution and logistics tasks and therefore play a key role i n subsequent benchmarking analysis.

Source: Organized by authors

Lower relative weights assigned to attributes such as responsibility or friendliness should not be interpreted as indicators of low importance. Rather, these characteristics reflect baseline expectations that are implicitly assumed by employers and therefore less frequently emphasized in vacancy descriptions compared to task-specific or technology-related competencies.

Whereas operational and routine functions previously occupied a central position in job requirements, employers now increasingly prioritize higher-order competencies of an analytical, managerial, and cross-functional nature. This contrast reflects not only a quantitative redistribution of demand but also a qualitative redefinition of professional profiles relevant for distribution-oriented activities. The interim industry-level results further indicate a gradual erosion of rigid boundaries between professional groups (Table 3).

Table 3: Results of Analysis of the Labor Market

Industry	Average salary (RUB)	Required competence	Dominant demand locations in Russia
IT	60579	Programming (Python, Java, C++), Data analysis, Cybersecurity, Cloud computing, Mobile app development	Moscow, St. Petersburg, Kazan, Novosibirsk
Engineering	51495	Mechanical engineering, Electronics, Automation, Design (AutoCAD, SolidWorks)	Urals, Volga region, large industrial regions
Economics	44812	Financial analysis, Accounting, Economic modeling, Risk management	Major cities, financial centers
Law	46532	Corporate law, practicing law, international law, financial law	Law firms, corporate sector in major cities

Source: Organized by authors

Competencies traditionally attributed to information technology or engineering specializations are increasingly required in trade, logistics, and distribution. This pattern confirms the emergence of hybrid professional profiles

characteristic of contemporary distribution systems, in which managerial, analytical, and customer service functions are structurally interrelated.

Application of the proposed methodology also enables the construction of graph-based representations describing relationships between competencies and vacancies.

Figure 5 illustrates such a model using the example of five vacancies in the information technology sector, demonstrating how individual competencies form interconnected structures rather than isolated skill elements.

Graph representations provide an additional analytical perspective on competence demand by revealing stable patterns of skill co-occurrence and the relational structure of competencies within specific occupational profiles. This representation supports the identification of hybrid competence configurations that are not readily captured through linear or frequency-based analytical approache

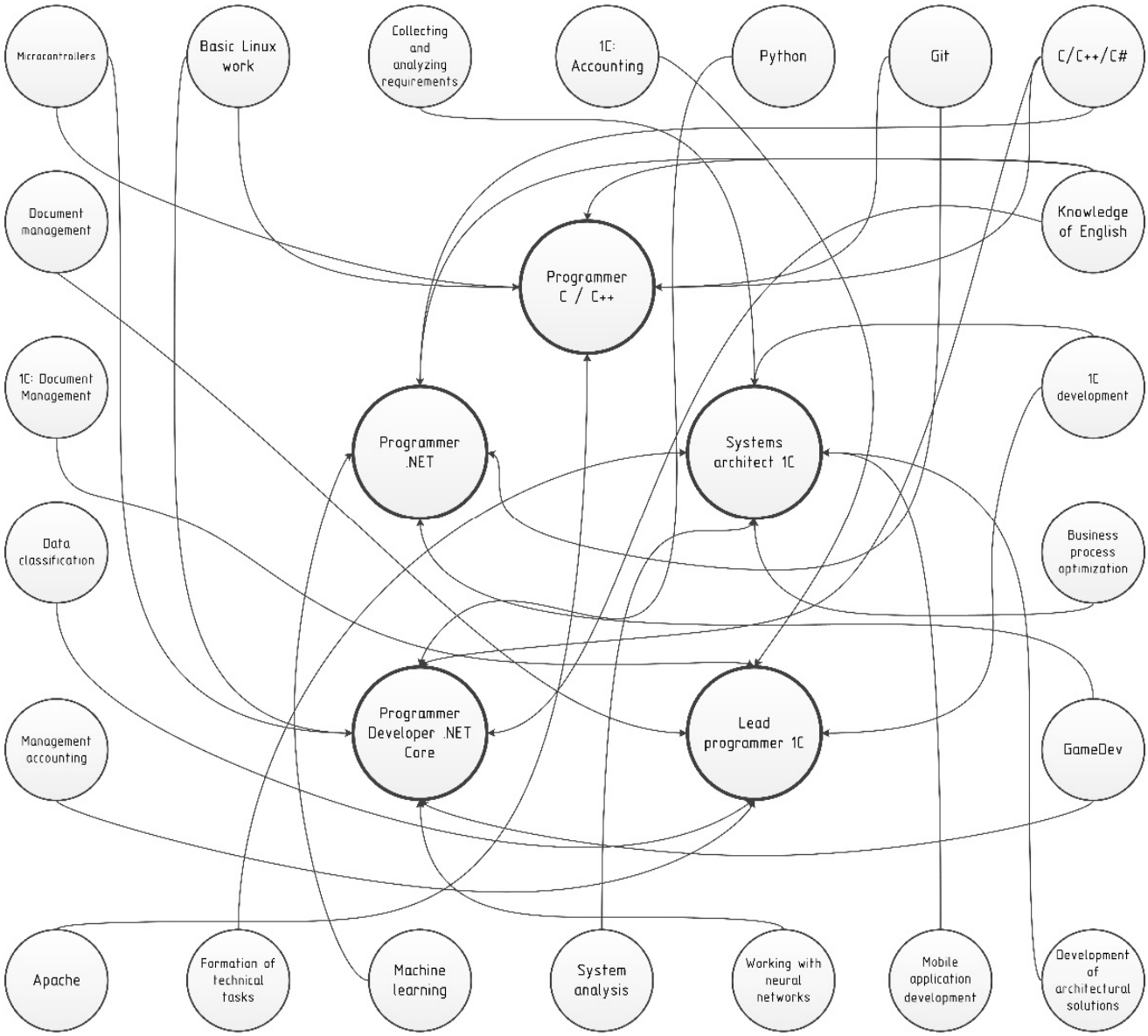

Figure 5: Graph Model Illustrating the Structure of Competencies Across Selected Vacancies

A key outcome of the pilot assessment is the identification of competencies exhibiting the most pronounced discrepancies between labor market requirements and the content of institutional competence representations (Table 4).

Table 4: Competencies with the Largest Gaps and Underestimated Skills in Institutional Competence Structures

Category	Competencies with the largest gap	Underestimated skills in education programs
Analytical skills	Data analysis, interpretation of complex reports	Critical thinking, statistical data processing
Digital competencies	Programming, working with big data, automation	Cybersecurity, digital literacy, AI and ML
Business and management	Financial analysis, strategic planning	Risk management, project management
Communication skills	Negotiation skills, public speaking	Personal branding, digital communication
Legal literacy	Interpretation of regulatory legal acts	International law, compliance
Engineering and technical skills	Technical modeling, working with high-precision systems	Automated control systems, additive technologies
Soft skills	Leadership, adaptability, emotional intelligence	Time management, stress resistance
Customer management	Customer orientation, personalized service	Customer expectation management, service design
Economics and finance	Budgeting, investment analysis	Financial literacy, fintech tools

Source: Organized by authors

These discrepancies primarily concern analytical and digital competencies, managerial capabilities, as well as skills related to customer interaction and service design.

From a distribution system perspective, the identified gaps are not merely inventory lists of missing skills but indicators of specific operational vulnerabilities. The underrepresentation of Data analysis and Statistical data processing (Analytical skills) points to a constrained capacity for predictive demand forecasting and inventory optimization, directly impacting stock turnover and warehousing efficiency. Gaps in Cybersecurity and AI and ML (Digital competencies) suggest systemic risks for digitally integrated supply chains and a lag in adopting automation for logistics coordination. The shortage of Risk management and Project management competencies (Business and management) can hinder the ability to manage supply chain disruptions and orchestrate complex digital transformation projects within distribution networks. Weak institutional reflection of Customer expectation management and Service design (Customer management) highlights a potential bottleneck in developing the sophisticated omnichannel service capabilities required for modern retail. Thus, systematic underrepresentation of these components signals more than a talent shortage. It reveals structural gaps in the competence infrastructure that supports the resilience, efficiency, and service quality of contemporary distribution systems.

4.2. Discussion of the Results in the Context of Distribution Systems

The empirical results compel a system-level interpretation. The identified competence discrepancies are not random noise in the labor market but reflect coherent, structural patterns that directly translate into constraints on distribution system performance. They are not limited to isolated or incidental mismatches but form clusters that map onto critical functional domains. For instance, the limited integration of data analytics and automation-related competencies signals a systemic barrier to improving inventory coordination and optimizing logistics processes, directly affecting key performance indicators such as carrying costs and order fulfillment cycle times. This interpretation is consistent with the observed concentration of vacancy requirements in digital/analytical skills and the recurring underrepresentation of these components in program descriptions (Table 4). At the same time, the relatively weak presence of managerial and customer-oriented competencies narrows the ability of organizations to maintain supply chain stability and to adjust operational practices in response to omnichannel interaction with consumers.

These results show that conventional labor market tools have limited explanatory power. Approaches relying on aggregated statistical indicators or expert-based assessments tend to reproduce established occupational categories. As a result, they are less effective in detecting gradual or latent shifts in competence demand associated with digital transformation and increasing organizational complexity in distribution networks. By contrast, the application of machine learning techniques to the semantic analysis of vacancy texts allows emerging and hybrid competence configurations to be identified at earlier stages, before they become formally embedded in occupational classifications or educational standards.

An additional implication of the results concerns the interaction between the labor market and the education and training system. The findings indicate the importance of moving toward a feedback-oriented model in which employer demand and educational content are linked through systematic and recurrent analytical procedures. Within this framework, competence structure adjustment development is informed by continuous evidence derived from labor market data. This approach moves beyond episodic revisions or declarative alignment with formal standards. Such an approach is particularly relevant for retail and distribution systems, where technological, organizational, and customer-related changes tend to occur

simultaneously and to reinforce one another. From an operational perspective, the identified competencies should be interpreted as interrelated bundles rather than as isolated requirements. As illustrated by the network structure in Figure 5, individual job roles are characterized by the co-occurrence of technical, analytical, and coordination-related skills, reflecting the integrated nature of contemporary job design.

In distribution and logistics contexts, similar hybrid configurations emerge, where digital and analytical competencies support inventory planning, demand forecasting, and coordination of logistics flows, while managerial competencies enable decision-making under uncertainty and alignment across supply chain actors. Customer-oriented competencies complement these functions by supporting omnichannel interaction and service processes. The combination of these competencies within single roles is directly associated with operational performance outcomes, including inventory turnover, service level, and supply chain responsiveness.

Taken together, the results of the study show that the proposed methodological approach is suitable for practical workforce development tasks in retail and distribution while remaining analytically robust.

The analysis demonstrates that the proposed diagnostic methodology can effectively map the evolving competence landscape of retail and distribution systems. The primary value of this map lies in its ability to inform data-driven strategies for workforce development and organizational adaptation within the sector. For educational institutions, these insights offer externally validated signals for strategic consideration, positioning institutional framework adaptation evolution as a responsive activity within a broader ecosystem focused on mitigating systemic competence bottlenecks and enhancing distribution network performance.

5. Conclusions

This study examines workforce competence mismatches as a structural constraint on retail and distribution systems, affecting their functioning and adaptation amid digital transformation and growing supply chain complexity. The results indicate that contemporary distribution systems increasingly depend on hybrid competence configurations. However, existing institutional mechanisms for reproducing these competencies fail to keep pace with technological and organizational change. From a distribution-system perspective, such mismatches represent not isolated educational deficiencies, but factors that limit operational coordination and overall system adaptability.

The analysis confirms the relevance of an integrated methodological approach combining vacancy-based labor market analytics, machine learning, and semantic modeling to diagnose workforce-related bottlenecks in retail and distribution. This approach detects structural competence mismatches that remain obscure in aggregated labor statistics or expert assessments. These mismatches are most pronounced in digital, analytical, managerial, and customer-oriented competence bundles. Systemically, they form a persistent constraint on logistics coordination, inventory management, and omnichannel service performance within distribution networks.

Empirical evidence from the Penza region demonstrates that competencies traditionally associated with information technology and engineering are increasingly embedded across a wide range of distribution-related occupations. This trend reflects the emergence of hybrid professional profiles that integrate analytical, managerial, and service-oriented functions within unified job roles. Within retail and distribution systems, such profiles are becoming a necessary condition for effective inventory control, logistics coordination, and the organization of omnichannel interactions with consumers. The Penza region serves in this study as a pilot case for testing the proposed analytical framework, rather than as a basis for generalizing absolute demand levels.

From an applied perspective, the proposed framework provides a diagnostic tool for identifying workforce constraints and coordination inefficiencies. Within this logic, educational programs are interpreted as a secondary, institutional domain that formalizes and reproduces competence demand generated by distribution systems. Consequently, analytically grounded adjustments to educational content represent a downstream application of the findings, not the study's primary focus.

The theoretical contribution of this work lies in conceptualizing workforce competence as a distribution-system resource and treating competence misalignment as a form of distribution inefficiency. By linking vacancy-based analytics to system-level interpretation, the study contributes to distribution science, demonstrating how labor market data can support adaptive coordination in retail and distribution networks undergoing digital transformation.

These findings have direct implications for management practice. For workforce planning, the presented methodology enables managers to monitor skill demand in real time and adjust recruitment strategies. Job profiles for critical functions, such as inventory management, should explicitly require blended skills in data analytics and logistics operations. To address systemic shortages, organizations may benefit from prioritizing investment in targeted training programs, particularly in areas like data analysis, cybersecurity, and omnichannel service design, to

strengthen their distribution networks. Collaboration with universities and training providers can help align education with industry needs. Sharing concrete evidence of competence gaps enables educational institutions to adapt curricula, creating a more responsive talent pipeline. To measure the impact of workforce development, managers can link competence profiles to operational performance metrics. Tracking indicators such as inventory turnover rates and order fulfillment times alongside skill deployment helps quantify the return on investment in human capital. In regions with scarce specific competencies, flexible strategies become crucial. These may include offering remote work for analytical roles, partnering with specialized firms, or establishing internal training centers to cultivate necessary hybrid skillsets within the organization.

The results of this study should be interpreted in light of several limitations. First, the empirical analysis relies on data from a single region (Penza), which reflects a specific regional labor market and institutional context; therefore, the findings do not claim global universality. Second, the assessment of competence demand is based on requirements articulated in vacancy texts, which may underrepresent informal daily tasks and implicit soft skills not consistently specified in job advertisements. Third, the analysis focuses on declared institutional competence structures and does not directly evaluate how competencies are implemented in actual organizational or educational practices.

These limitations suggest directions for further research oriented toward distribution systems, including comparative cross-regional analysis, examination of competence demand across different distribution formats (retail chains, distribution centers, e-commerce platforms), and the development of longitudinal monitoring frameworks to assess how workforce competence dynamics influence the resilience and performance of distribution systems over time.

Acknowledgements

This study has been funded by the Science Committee MSHE RK “Modernization of the quality assurance system of higher education in Kazakhstan based on digitalization: development of approaches, mechanisms and information base” BR24992974.

References

Assante, D., Caforio, A., Flamini, M., & Romano, E. (2019). Smart Education in the context of Industry 4.0. *2019 IEEE Global Engineering Education Conference (EDUCON)*, 1140-1145.

Bakhshi, H., Downing, J. M., Osborne, M. A., & Schneider, P. (2017). *The Future of Skills: Employment in 2030.* London: Pearson.

Beresewicz, M., Wydmuch, M., Cherniaiev, H., & Pater, R. (2025). Multilingual Hierarchical Classification of Job Advertisements for Job Vacancy Statistics. *Journal of Official Statistics*, *0*(0). https://doi.org/10.1177/0282423X251395400

Boselli, R., Cesarini, M., Mercorio, F., & Mezzanzanica, M. (2018). Classifying online job advertisements through machine learning. *Future Generation Computer Systems, 86, 319-328.* https://doi.org/10.1016/j.future.2018.03.035

Botov, D., Klenin, J., Melnikov, A., Dmitrin, Y., Nikolaev, I., Vinel, M. (2019). Mining Labor Market Requirements Using Distributional Semantic Models and Deep Learning. In: Abramowicz, W., Corchuelo, R. (eds) *Business Information Systems.* BIS 2019. Lecture Notes in Business Information Processing, vol 354. Springer, Cham. https://doi.org/10.1007/978-3-030-20482-2_15

Chen, J., Li, K., Liu, Z., Zhang, T., Wen, W., Song, Z., Huang, T. (2019). Data analysis and knowledge discovery in web recruitment—Based on big data related jobs. *2019 International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), 142-146.* https://doi.org/10.1109/mlbdbi48998.2019.00033

Coolen, P., van den Heuvel, S., van de Voorde, K., & Paauwe, J. (2023). Understanding the adoption and institutionalization of workforce analytics: A systematic literature review and research agenda. *Human Resource Management Review, 33*(4), 100985. https://doi.org/10.1016/j.hrmr.2023.100985

Dadzie, A.-S., Sibarani, E. M., Novalija, I., & Scerri, S. (2018). Structuring visual exploratory analysis of skill demand. *Journal of Web Semantics, 49*, 51–70. https://doi.org/10.1016/j.websem.2017.12.004

Deev, M., Gamidullaeva, L., Finogeev, A., & Vasin, S. (2021). The convergence model of education for sustainability in the transition to the digital economy. *Sustainability, 13(20),* 11441. https://doi.org/10.3390/su132011441

Deming, D. J., & Noray, K. (2020). Earnings dynamics, changing job skills, and STEM careers. *The Quarterly Journal of Economics, 135*(4), 1965–2005. https://doi.org/10.1093/qje/qjaa021

European Commission: Eurostat, *Inferring job vacancies from online job advertisements – 2021 edition*, Publications Office, 2021, https://data.europa.eu/doi/10.2785/963837

Grönroos, C. & Voima, F.P. (2013) Critical Service Logic: Making Sense of Value Creation and Co-Creation. *Journal of the Academy of Marketing Science, 41*, 133-150. https://doi.org/10.1007/s11747-012-0308-3

Jiang, W., Feng, T., & Wong, C. Y. (2025). How big data analytics capabilities drive supply chain resilience: The mediating roles of supply chain visibility and flexibility. *International Journal of Physical Distribution & Logistics Management*, *55*(5), 505–539. https://doi.org/10.1108/IJPDLM-07-2024-0252

Jing, H., & Fan, Y. (2024). Digital transformation, supply chain integration and supply chain performance: Evidence from Chinese manufacturing listed firms. *SAGE Open, 14*(3). https://doi.org/10.1177/21582440241281616

Khaouja, I., Kassou, I., & Ghogho, M. (2021). *A survey on skill identification from online job ads*. IEEE Access, 9, 118134–118153. https://doi.org/10.1109/ACCESS.2021.3106120

Kim, J. W., Rhee, J. H., & Park, C. H. (2024). How Does Digital Transformation Improve Supply Chain Performance: A

Manufacturer's Perspective. *Sustainability*, *16*(7), 3046. https://doi.org/10.3390/su16073046

Kitcharoen, K. (2023). The adoption of Big Data to achieve firm performance of global logistic companies in Thailand. *Journal of Distribution Science, 21*(1), 53–63. https://doi.org/10.15722/jds.21.01.202301.53

Kureková, L. M., Beblavý, M., & Thum-Thysen, A. (2015). Using online vacancies and web surveys to analyse the labour market: A methodological inquiry. *IZA Journal of Labor Economics, 4*(1).-1-20, https://doi.org/10.1186/s40172-015-0034-4

Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. In *Proceedings of the 31st International Conference on Machine Learning (ICML)*, 1188–1196. https://doi.org/10.48550/arXiv.1405.4053

McGuinness, S., Pouliakas, K., & Redmond, P. (2018). Skills mismatch: Concepts, measurement and policy approaches. *Journal of Economic Surveys, 32*(4), 985–1015. https://doi.org/10.1111/joes.12254

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *Transactions of the Association for Computational Linguistics, 1*, 1–12. https://doi.org/10.48550/arXiv.1301.3781

Miranda, S., Della Noia, T., Ostuni, V. C., & Zanker, M. (2017). An ontology-based model for competence management. *Data & Knowledge Engineering, 107*, 51–76. https://doi.org/10.1016/j.datak.2016.12.001

Mkrttchian, V., Gamidullaeva, L., Finogeev, A., Chernyshenko, S., Chernyshenko, V., Amirov, D., & Potapova, I. (2021). Big data and Internet of Things (IoT) technologies' influence on higher education: Current state and future prospects. *International Journal of Web-Based Learning and Teaching Technologies, 16*(5), 137–157. https://doi.org/10.4018/IJWLTT.20210901.oa8

Mkrttchian, V., Krevskiy, I., Bershadsky, A., Glotova, T., Gamidullaeva, L., & Vasin, S. (2019). Web-based learning and development of university's electronic informational educational environment. *International Journal of Web-Based Learning and Teaching Technologies, 14*(1), 32–53. https://doi.org/10.4018/IJWLTT.2019010103

Mocanu, C., Zamfir, A. M., & Pirciog, S. (2014). Matching curricula with labour market needs for higher education: State of art, obstacles and facilitating factors. *Procedia - Social and Behavioral Sciences, 149*, 602–606. https://doi.org/10.1016/j.sbspro.2014.08.234

OECD. (2017). *OECD Skills Outlook 2017: Skills and Global Value Chains.* Paris: OECD Publishing. https://doi.org/10.1787/9789264273351-en.

OECD. (2021). *OECD Employment Outlook 2021: Navigating the COVID-19 crisis and recovery*. OECD Publishing. https://doi.org/10.1787/5a700c4b-en OECD

Pejić Bach, M., Bertoncel, T., Meško, M., & Krstić, Ž. (2020). Text mining of Industry 4.0 job advertisements. *International Journal of Information Management, 50*, 416–431. https://doi.org/10.1016/j.ijinfomgt.2019.07.014

Senger, E., Zhang, M., van der Goot, R., & Plank, B. (2024). *Deep learning-based computational job market analysis: A survey on skill extraction and classification from job postings*. In Proceedings of NLP4HR 2024 (pp. 1–15). Association for Computational Linguistics. https://aclanthology.org/2024.nlp4hr-1.1/

Sibarani, E. M., Scerri, S., Morales, C., Auer, S., & Collarana, D. (2017). Ontology-guided job market demand analysis: A cross-sectional study for the data science field. In *Proceedings of SEMANTiCS 2017* (pp. 25–32). https://doi.org/10.1145/3132218.3132228

Singh, R. K. (2025). Measuring supply chain resilience performance: role of data analytics, collaboration and flexibility. *Measuring Business Excellence, 29*(1), 121–136. https://doi.org/10.1108/MBE-01-2024-0007

Spada, I., Fabbroni, V., Chiarello, F., & Fantoni, G. (2024). Standardising job descriptions in the humanitarian supply chain: A text mining approach for recruitment process. *PLOS ONE, 19*(7), e0305961. https://doi.org/10.1371/journal.pone.0305961

Tavakoli, M., Faraji, A., Vrolijk, J., Molavi, M., Mol, S. T., & Kismihók, G. (2022). An AI-based open recommender system for personalized labor market–driven education. *Advanced Engineering Informatics, 52*, 101508. https://doi.org/10.1016/j.aei.2021.101508

Tolstykh, T., Gamidullaeva, L., & Shmeleva, N. (2021). Universities as knowledge integrators and cross-industry ecosystems: A self-organizational perspective. *SAGE Open*, *11*(1), 1–15. https://doi.org/10.1177/2158244020988704

Trenerry, B., Chng, S., Wang, Y., Suhaila, Z.S., Lim, S.S., Lu, H.Y., & Oh, P.H. (2021). Preparing workplaces for digital transformation: An integrative review and framework of multi-level factors. *Frontiers in Psychology, 12*, 620766. https://doi.org/10.3389/fpsyg.2021.620766

Tzimas, G., Zotos, N., Mourelatos, E., Giotopoulos, K. C., & Zervas, P. (2024). From Data to Insight: Transforming Online Job Postings into Labor-Market Intelligence. *Information*, *15*(8), 496. https://doi.org/10.3390/info15080496

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics, 34*(2), 77–84. https://doi.org/10.1111/jbl.12010

Appendix A. Technical Implementation Details of Data Collection and Competence Modeling

This appendix presents supplementary technical information related to the collection, preprocessing, and computational implementation of the competence analysis procedures applied in the study. The purpose of this appendix is to ensure methodological transparency and reproducibility while keeping the main text focused on analytical logic and economic interpretation.

Vacancy data used in the study were obtained through the Rosnavyk labor market analytics service, which integrates data from several major online recruitment platforms, including HeadHunter, SuperJob, and others. The dataset covers approximately 1.200 vacancy postings published in 2024 and reflects employer demand in retail, distribution, and logistics occupations within the regional labor market.

Job vacancy data were collected from open online sources using an automated system designed to extract textual descriptions of employer requirements. The system aggregates vacancy information from multiple platforms and consolidates it into a unified dataset suitable for further analysis. To ensure consistency of employer information and eliminate duplicate vacancy descriptions published across different sources, vacancy records were linked to spatial coordinates and timestamps. Address-based geocoding was applied using the Yandex. Map API, which made it possible to identify and merge records referring to the same employer and location.

Textual vacancy descriptions were cleaned and normalized to remove non-informative elements and formatting artifacts. Linguistic preprocessing was implemented using rule-based and dictionary-driven parsing procedures. In particular, the Tomita-parser was employed to extract structured information from unstructured vacancy texts on the basis of predefined grammar rules and dynamically updated keyword dictionaries. The output of this stage consisted of normalized textual representations suitable for semantic analysis.

Formalization of competence requirements was based on transforming vacancy texts and educational program descriptions into vector representations within a common semantic space. Keyword dictionaries constructed from both data sources were used to identify competence-related terms, which were subsequently represented in vector form to capture semantic relationships between skills an functional requirements. Model quality and coherence of competence groupings were assessed using standard similarity measures.

Clustering and benchmarking of competence vectors were implemented using neural network–based methods appropriate for high-dimensional textual data. These methods were applied to group competence profiles derived from employer vacancies and educational programs and to assess their relative alignment. The computational implementation relied on standard machine learning frameworks, including TensorFlow and Keras, as well as text processing libraries such as Gensim. Model configuration and tuning were oriented toward stability and interpretability of results rather than toward maximizing predictive accuracy.

Quantitative criteria used for assessing similarity between competence profiles and for classifying educational programs according to their degree of alignment with labor market demand are formally defined in Appendix B.

Appendix B. Formal Representation of the Competence Modeling Procedure

The most common method for measuring the semantic similarity between word vectors is by calculating the cosine similarity:

$$S(A,B) = cos(\theta) = \frac{\sum_{i=1}^{K} A_i B_i}{\sqrt{\sum_{i=1}^{K} A_i^2}\sqrt{\sum_{i=1}^{K} B_i^2}} \qquad (1)$$

In our case, we calculate the maximum cosine similarity between a vector and a nearby centroid, and the minimum cosine similarity between a vector and a distant centroid.

While the use of fixed reference centroids enhances interpretability of the results, it may limit the detection of unconventional or weakly institutionalized hybrid competence profiles, which can be addressed in future research through adaptive or dynamic clustering strategies.

The transformer's internal attention mechanisms assign weights to words in the input sequence based on their similarity to centroids in a clustering task.

These weights are used to adjust the representations of each word in the vector model. The sum of these weighted representations is used to calculate a normalized exponential softmax activation function at the output of the LSTM layer in the model. This activation function probabilistically ranks the similarity of the vector models to cluster centroids:

$$Softmax_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_i}}, \qquad (2)$$

where z_i are the embeddings of the input vector z, K is the number of embeddings.

Three vectors are created for each word in the competence vector: a query, a key, and a value. These

vectors help determine the significance of each word in the competence description with respect to other words in the sentence. This allows the model to identify which words are most important in the description based on their context and relationship with other words.

The clustering algorithm that uses the calculation of semantic similarity between vector competence models works as follows:

The number of clusters of competence vector T is set and the vector representations (embeddings) are chosen as centroids depending on the direct or inverse clustering problem.

The input sets of competence vectors are represented as an embedding matrix Xj (j=1, ..., N). Each embedding defines a point in the multidimensional clustering space with previously defined centroids CL(k) (k=1, ..., T). The cosine similarity function is $S_{A_iB_j}^{(k)}$.

1. The coordinates of the centroids in the CL^k space *(k=1, ..., T)* we calculated as: $CL_k = \frac{\Sigma_{k=1}^{T}(\mu_{X_j}^{(k)} X_j)}{\Sigma_{k=1}^{T}(\mu_{X_j}^{(k)})}$ (3)

2. The cosine similarity function is used to calculate the semantic similarity between embeddings and centroids $S_{A_iB_j}^{(k)}$. Similarity thresholds (high, medium, low) were introduced to support analytical interpretation of alignment between vacancy-based competence profiles and educational programs rather than strict binary classification.

3. The similarity matrix helps to calculate the integral function of the cosine similarity between all embeddings and centroids:

$$\mu_{X_j}^{(k)} = \frac{1}{\Sigma_{k=1}^{T}\left(\frac{RS_{X_i}^{(k)}}{RS_{X_j}^{(k)}}\right)^{2/s-1}}), \quad (4)$$

4. To stop the iterative process of recalculating the matrix, a parameter ε > 0 is set. If the condition $\{|\mu_{X_j}^{(k)} - \mu_{X_j}^{(k-1)}|\} < \varepsilon$ is not met, we return to the fifth step.