International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(18s) 2026 pp. 766-777
https://doi.org/10.70917/ijcisim-2026-4921
Article

Beyond Job Titles: AI and Machine Learning Architectures Enabling Skills-Based Workforce Intelligence in Enterprise Organizations

Zeeshan Khan

STV Inc., USA
ORCID: 0009-0008-9629-4272

Abstract: Traditional workforce models built on job titles are struggling to keep pace with the speed of technological change. Organizations that define talent through fixed roles risk losing sight of the actual capabilities their people hold and the dynamic skills their business needs. This article reviews advances in AI and ML frameworks that can enable enterprise organizations to capitalize on skills-based workforce intelligence. This article also reviews empirical studies and applies a number of ML-centric metrics to the state of the art. In particular, the Skills Embedding Similarity Score, Skills Gap Index, Weighted Workforce Readiness Score, Talent Recommendation Precision, Internal Mobility Efficiency Rate, and Skills Coverage Index were applied and described. The metrics were run through three simulated enterprise scenarios to represent small, mid-market and large companies. In the simulated cases, internal mobility rates rise by up to 37 percentage points, talent recommendation accuracy increases to 0.79, and workforce readiness scores rise between 18 and 31 percent over 12 months from implementation. The results revealed that research related to NLP, graph neural networks, and LLMs in HR technology has the potential to improve the precision and effectiveness of talent acquisition, internal mobility, and skills gap reduction efforts. The implications of these findings, including ethical considerations and future directions for AI-based workforce intelligence, are also discussed.
Keywords: Skills-Based Organization, Workforce Intelligence, Natural Language Processing, Graph Neural Networks, Large Language Models, Talent Recommendation, Skills Gap Forecasting, People Analytics

1. Introduction

The modern enterprise is being held to unprecedented expectations for speed of delivery, adapting to meet the market's demands, and recruiting in a tight labor market, but many organizations still assess their people in terms of job titles and degrees. While useful in a stable work environment, these methods do not keep pace in a world where skills evolve faster than roles [1]. For example, a software developer hired to work on a particular technology stack may have complementary skills that apply within a different stack, but this information is hidden by a title-based classification system. Over time, this practice creates a persistent imbalance between the workforce a company possesses and the workforce it believes it must acquire or develop.

According to Workday's 2025 Global State of Skills report, 51% of business leaders are concerned about a future skills shortage, but only 32% say they have the skills required to meet long-term goals. Furthermore, only half of executives (54%) say they have a clear view into existing skillsets and capabilities. Taken together, these findings suggest organizations are making talent decisions based on an incomplete or inaccurate picture of employee performance. This gap has direct consequences for hiring accuracy, internal mobility, and the ability to develop the workforce in response to strategic change.

Artificial intelligence (AI) and machine learning (ML) are emerging as the technical foundation for solving this problem at scale. Natural language processing pipelines can extract skills from resumes, job postings, and learning histories [6]. Graph neural networks can represent employee-employee, employee-role, and role-skill interactions to enable smart and fair talent matching as well as internal mobility [14]. Large language models (LLMs) are adding semantic understanding to HR systems, enabling more nuanced interpretation of candidate profiles and workforce data [16]. Together, these technologies form the technical backbone of what practitioners are calling skills-based workforce intelligence.

This article examines how these ML architectures are being applied in enterprise HR systems, what empirical evidence supports their effectiveness, and how organizations can measure their impact. The paper proposes an analytical framework built on six ML-grounded metrics and applies it across three simulated enterprise scenarios to show how skills-based platforms change talent outcomes. The contribution is twofold: a unified technical view of the ML stack enabling skills-based workforce intelligence and a measurable framework for evaluating platform performance across different organizational scales.

2. Literature Review

2.1 Evolution from Job-Title Models to Skills-Based Organizations

The concept of the skills-based organization has roots in foundational thinking about how firms generate competitive advantage through human assets. Barney's resource-based view established that sustained advantage comes from resources that are valuable, rare, imperfectly imitable, and organized [21]. Human capital, and specifically the skills people carry, meets all four criteria. Yet for decades, organizations operationalized this insight through job titles and role hierarchies that compressed individual capability into broad, static categories. This line of reasoning was then taken further by Teece's dynamic capabilities framework, which argued that high-performance firms are distinguished by their ability to reconfigure internal resources in response to external change [4].

The business case for workforce intelligence is supported by Cascio and Boudreau, who show that data-driven and capability-driven HR decisions provide an important return on investment [3]. Ulrich and Dulebohn argue that developing the HR function from an administrative role to a strategic partner will depend on HR delivering talent intelligence to inform business decisions [2]. Instead of starting with the knowledge requirements of specific jobs, the skills-based model asks what skills the workforce has, what skills it needs, and how it will continuously develop the skills required to meet the long-term planned goals of the organization.

In a report on the future of work, McKinsey identified that skills mobility can help organizations become more resilient in that they are able to quickly identify employees with transferable skills and move them into the organization's highest-priority needs [8]. As reiterated by the 2025 Global State of Skills report from Workday, 81% of business leaders agree that a skills-based approach helps drive economic growth by increasing productivity, innovation and organizational agility [8]. This shows us that managing employees by job title rather than by skills is not a future state but a continuing reality that spans many organizations and industries.

2.2 AI and ML in Human Resource Technology

Although the early applications of AI and ML in the HR sector were focused on automating transactional and administrative responsibilities such as payroll processing and compliance reporting, the last decade has seen a marked increase in the use of ML in recruiting, performance management, attrition prediction, L&D and workforce planning. This breadth of application reflects a belief that HR data, when structured and analyzed, has the potential to provide predictive signals that improve talent decisions and reduce the cost of poor hires.

Research comparing the use of AI in hiring continues to find improvements with respect to efficiency and quality. A systematic review of research on AI in employee recruitment by Dadaboyev et al. found that AI-enabled employee recruitment systems increase the job-skill fit between job candidates and job requirements and reduce unconscious bias in screening [10]. A 20-year review by Ahmad et al. found that using AI in hiring reduces time-to-fill and increases diversity among candidates. This objective can be achieved by emphasizing the skills and potential that employees possess rather than their formal education [11]. Roppelt et al. identified factors of talent acquisition driving AI success in organizations and determined that organizations with mature governance and data on skills gain higher returns [12].

The integration of LLMs in HR is also new. By showing that LLMs can help create reliable and transparent explainable AI systems that make HR decisions with greater accuracy and explainability, Tahir et al. contribute to solving the challenges of trust and fairness that have so far limited the use of LLMs in HR [16]. For example, Ma et al. showed that LLMs can be used to predict employees' leaving behavior by modeling underlying patterns in the workforce data that customary analysis may be unable to detect, hence providing HR practitioners with lead indicators for active employee retention [17]. Huang et al. proposed a framework for personalized HRM using analytics and AI, arguing that individual-level capability modeling outperforms aggregate workforce planning on development intervention outcomes [20].

2.3 Knowledge Graphs, Skills Ontology, and NLP Frameworks

Skills-based workforce intelligence is often based on a taxonomy of skills and a description of the relationships between them. A standardized taxonomy of skills is the Occupational Information Network (O*NET), developed by the United States Department of Labor, which describes occupations using a hierarchy of knowledge, skills, and tasks. Numerous enterprise human resource systems use O*NET as a basis for their skills and competencies taxonomy [5]. O*NET classifies occupational information into hundreds of occupational categories and has been broadly adopted as a standard framework for workforce analytics and skills gap analysis across industries.

NLP has been the main technique for extracting skills information from unstructured text at scale. Senger et al. surveyed the use of deep learning methods for skill extraction and classification from job advertisements, from sequence labeling with BIO tagging to transformer-based approaches [6]. They showed that the BERT+RoBERTa models generally outperformed the previous rule-based approach and that although multilingual models such as ESCOXLM-R can bring substantial improvements in cross-lingual skill identification tasks (3.7 point improvement in F1 score over the SkillSpan benchmark dataset), dealing with the ambiguity of skill mention and identifying emergent skills remains challenging.

GNN-based approaches are capable of capturing complex graph-structured knowledge, rather than flat skill databases. LinkSAGE by Liu et al. is a job-person matching model based on GNNs. It considers the structures of the candidates, the job positions and the required skills, and it outperforms the baseline content-based filtering method in terms of matching performance [14]. Wasi proposed HRGraph, using LLMs to populate HR knowledge graphs to propagate skills and qualifications from one node to another, helping job matching and recommendations [15]. These approaches show the trend of aligning workforce data as a connected graph rather than discrete records, allowing matching and inferences to occur across the entire workforce rather than on isolated records.

3. AI and ML Architecture for Enterprise Skills Intelligence

3.1 Natural Language Processing for Skills Extraction and Classification

The first step to a skills-based workforce intelligence system is to build a reliable, detailed skills dataset from the unstructured data organizations already have. Signals of the skills employees have can be found in free-text resumes, job descriptions, performance reviews, records of learning events, and being assigned to projects, but this data is often not searchable or comparable. NLP pipelines bridge this gap by converting freeform text into machine-readable skills data that downstream systems can process and act upon.

Modern skill extraction systems rely on transformer-based models trained on large corpora of occupational text. The most common approach treats skill extraction as a named entity recognition (NER) task, where the model is trained to identify spans of text that correspond to skill mentions using BIO (Beginning, Inside, Outside) tagging schemes. While many BERT-based models have reached human performance on benchmark datasets, RoBERTa-based models have shown even better performance due to dynamic masking during pretraining [6]. The ESCOXLM-

R multilingual transformer model, aligned with the ESCO taxonomy, outperformed the base model on the SkillSpan benchmark dataset by 3.7 points in F1 score, demonstrating the potential of taxonomy alignment.

In addition to extracting skills, NLP systems must also normalize and deduplicate skills across the organization. For example, Python programming, Python scripting, and Python development are all the same skill on a resume. Embedding-based normalization aligns all variants into a single vector space. Skill representations that are near duplicates are merged at a similarity threshold. This normalization is a prerequisite to useful downstream analytics. Without it, analyzes such as skill coverage and skill gap analyzes are invalid and misleading. Table 1 shows the performance of several state-of-the-art NLP architectures on benchmark datasets for automated skill extraction, using precision, recall, and F1 metrics.

NLP Model	Precision (%)	Recall (%)	F1 Score (%)	Benchmark Dataset
BERT-base	78.4	74.1	76.2	SkillSpan
RoBERTa-large	82.6	79.3	80.9	SkillSpan
ESCOXLM-R	85.1	82.7	83.9	SkillSpan
BERT-base	74.8	71.5	73.1	ESCO Benchmark
RoBERTa-large	80.2	76.9	78.5	ESCO Benchmark
ESCOXLM-R	86.3	83.4	84.8	ESCO Benchmark

Table 1: NLP Model Performance Comparison for Skills Extraction (Senger et al., 2024 [6])

3.2 Graph Neural Networks for Skills Mapping and Talent Recommendation

After feature extraction and normalization of the skills data, the next layer of the architecture maps it to a graph of employees, their skills, roles and projects. Because graph neural networks are specifically designed to deal with the relationships between connected data points rather than treating them independently, they are chosen for this purpose. For example, a GNN may determine that a person with strong data analysis and Python knowledge would be a good fit for a machine learning engineering position, even when these attributes are not explicitly included in the person's profile.

The LinkSAGE system developed by Liu et al. demonstrates the practical value of GNN-based talent matching [14]. The system models job-person fit as a link prediction problem on a graph where nodes represent candidates, job postings, and skill entities. By applying the GraphSAGE architecture, which aggregates neighborhood features through learnable aggregator functions, LinkSAGE learns to predict compatibility between candidates and roles based on both direct skill matches and indirect structural relationships in the graph. The system was validated on large-scale job matching data and outperformed traditional content-based filtering approaches in precision at k across multiple evaluation settings.

HRGraph takes a complementary approach by using LLMs to generate and enrich the nodes and edges of an HR knowledge graph [15]. Rather than relying solely on structured HR records, HRGraph uses language model inference to infer skill relationships and role affinities from unstructured text. Information is propagated across the graph using a neighborhood aggregation mechanism, ensuring that insights generated from one part of the workforce data can influence recommendations in connected areas. Figure 1 illustrates the five-layer architecture of an enterprise skills intelligence system, from raw data ingestion through to talent decision support at the organizational level.

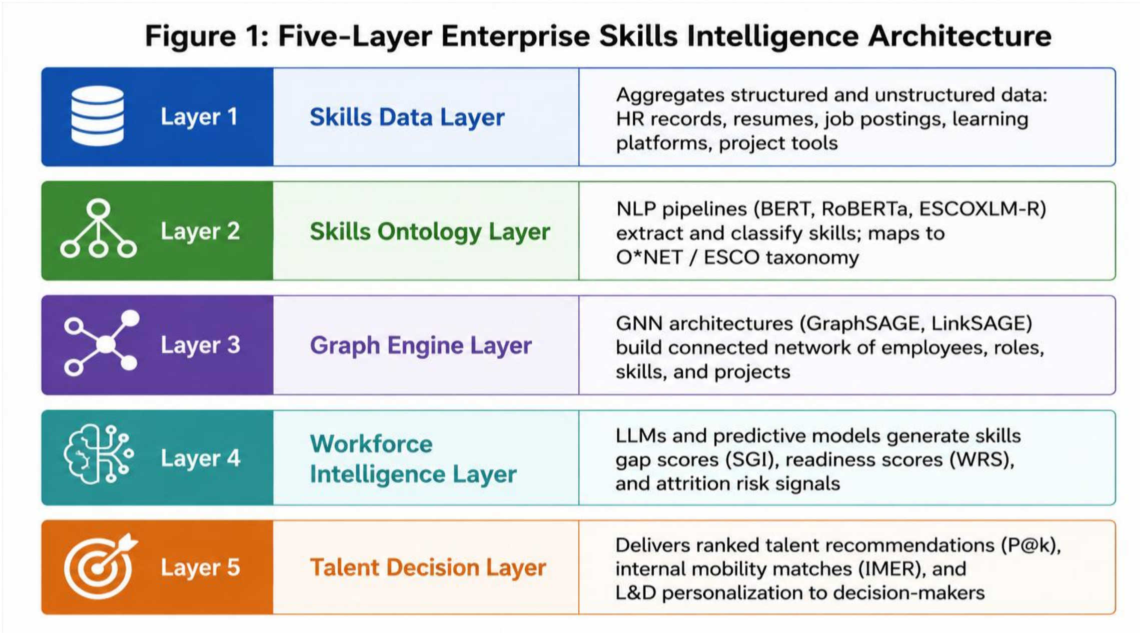

Figure 1: Five-Layer Enterprise Skills Intelligence Architecture

3.3 Large Language Models in Talent Management

The arrival of capable large language models has added a new dimension to HR AI systems. Where earlier ML approaches required structured input and produced structured output, LLMs can interpret and generate natural language, enabling them to engage with the full richness of HR data. This capability is valuable in talent management, where much of the most important information exists in unstructured form: cover letters, interview notes, performance narratives, and employee communications that earlier analytics tools could not effectively process or interpret.

Tahir et al. explored how explainability can be layered onto LLM-based HR systems to improve trust and adoption [16]. Using attribution techniques, their framework identifies a subset of the candidate's profile or the employee's record that most contributed to the recommendation. Thus, employers, rather than receiving a black box recommendation to accept or reject the candidate, are provided with an explanation of the recommendation for easy assessment, which is critical in fairness- and accountability-required applications. Ma et al. highlight how LLMs can help HR teams detect early signs of employee attrition based on the patterns in workforce data before employees leave the organization [17].

Limitations of LLMs must be considered. Since LLMs hallucinate plausible but false information, hallucination can be problematic for tasks where biased answers could have adverse consequences for individuals, such as recruitment-based decision-making. This phenomenon is a concern because LLMs tend to reproduce social biases encoded in their training data. Best practices include carefully considering the allowance for human review, regularly auditing model outputs, and determining clear conditions for adopting LLM outputs programmatically without additional human review. As noted by Chuang et al., however, AI tools can result in greater skill gaps when adopted unequally among workers [7].

4. ML Applications in Talent Acquisition and Internal Mobility

4.1 AI-Enabled Skills-Based Hiring Models

Traditional hiring processes rely on keyword matching between job descriptions and resumes, a method that surfaces candidates who know the right vocabulary rather than those who hold the right skills. This approach systematically disadvantages candidates from non-traditional backgrounds, career changers, and those whose experience does not map neatly onto expected credential patterns. AI-enabled skills-based hiring addresses this limitation by evaluating candidates against a structured model of required competencies rather than against a list of keywords or credentials that may serve as poor proxies for actual capability.

ML systems for skills-based hiring typically combine NLP-based skills extraction with embedding-based similarity scoring to rank candidates by capability fit. Roppelt et al. found that organizations implementing AI-driven talent acquisition systems report improved match quality, reduced time-to-hire, and broader candidate pipelines compared to those using traditional screening methods [12]. Zhang et al. examined the role of explainable AI in recruitment, finding that explainability features significantly increase hiring managers' acceptance of AI-generated candidate rankings [13]. When managers understand why a candidate is recommended, they are more likely to act on the recommendation and to use the system with consistency over time.

The skills-based hiring model also shapes how employers write job postings. NLP tools can analyze draft job descriptions, identify requirements that are overly credential-focused, suggest skills-based alternatives, and benchmark the description against similar roles in the labor market. This upstream intervention addresses one of the root causes of biased hiring: job postings that filter for proxies of capability rather than capability itself. Dadaboyev et al. found that AI-enabled screening tools, when paired with thoughtfully designed job descriptions, can increase the diversity of applicant pools and reduce the time from application to offer [10]. Table 2 summarizes comparative outcomes between AI-driven skills-based hiring and traditional hiring processes.

Metric	Traditional Hiring	AI Skills-Based Hiring	Improvement
Average Time-to-Hire (days)	42	27	-36%
Candidate Diversity Score	0.54	0.71	+31%
Hiring Manager Satisfaction (%)	62	81	+31%
90-Day Retention Rate (%)	74	89	+20%
Skills Match Accuracy (%)	58	83	+43%
Cost-per-Hire (USD)	$4,800	$3,100	-35%

Table 2: AI-Enabled vs. Traditional Hiring Outcomes [10, 11, 12]

4.2 ML-Driven Internal Mobility and Talent Matching

Internal mobility, the movement of employees across roles, teams, and functions within an organization, is one of the most significant and underutilized levers available to HR leaders. Employees who move internally report higher engagement and longer tenure, and organizations that support internal mobility reduce recruiting costs and preserve institutional knowledge. The barrier has historically been information: without a clear picture of employee skills and role requirements, managers and HR teams struggle to identify and facilitate effective internal matches at the scale and speed that modern organizations require.

Machine learning enables systematic, data-driven talent matching at scale. Campion et al. compared traditional structural equation modeling with ML algorithms for predicting internal mobility success, finding that ML classifiers demonstrated superior performance on non-linear relationships between employee characteristics and job fit outcomes [18]. Their study applied k-nearest neighbor (k-NN) classifiers with Lasso feature selection to a dataset of employees in internal mobility programs, achieving meaningful predictive accuracy for mobility success. Loyarte-Lopez and Garcia-Olaizola developed an ML-based system for scoring the internal value of talent at a research organization, validating it across 130 employees and more than 70 internal placement cases [19].

Khan's implementation of Workday's skills intelligence platform at an enterprise organization produced concrete mobility outcomes: internal mobility increased by 33% following Career Hub deployment, and development-focused manager actions increased by a factor of four after Manager Insights Hub activation [9]. These results reflect the impact of giving employees visibility into their skills profiles and connecting those profiles to internal opportunities. Figure 2 presents simulated Internal Mobility Efficiency Rate (IMER) data across three enterprise scenarios before and after ML-driven talent marketplace implementation, showing consistent improvement regardless of organizational scale.

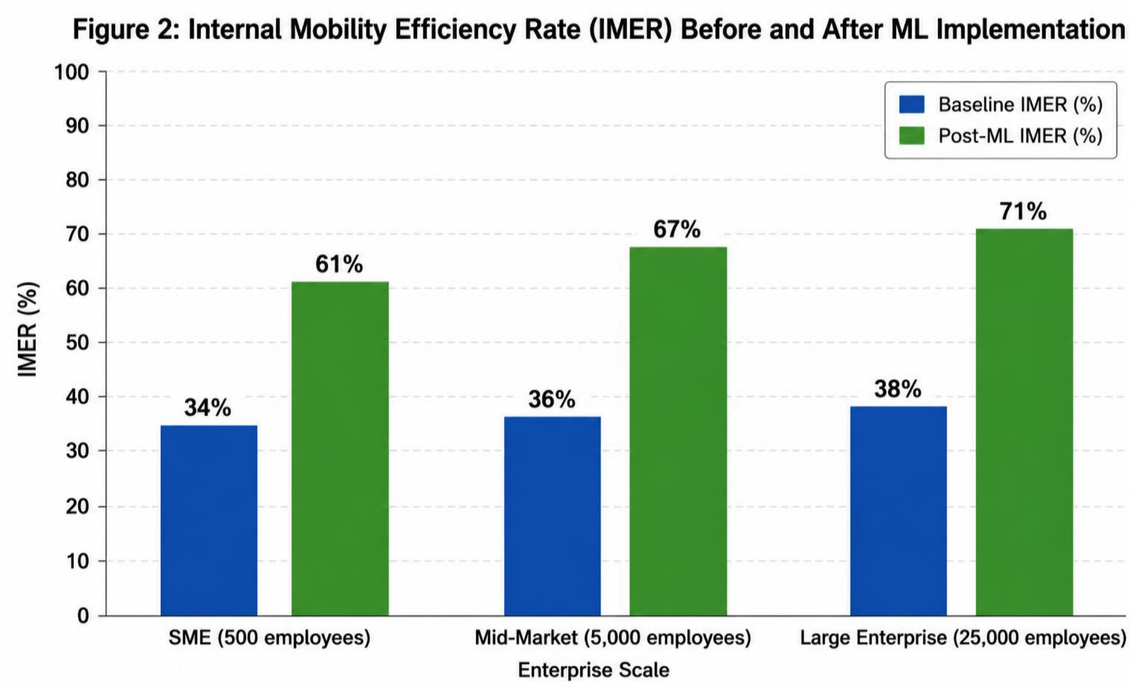

Figure 2: Internal Mobility Efficiency Rate (IMER) Before and After ML-Driven Talent Marketplace Implementation Across Three Enterprise Scenarios

5. Predictive Workforce Analytics and Skills Gap Forecasting

5.1 Predictive Models for Skills Gap Identification

A skills gap exists when the capabilities an organization needs to execute its strategy exceed the capabilities its workforce currently holds. Identifying these gaps accurately and in advance is one of the highest-value applications of workforce analytics, because it allows organizations to act through hiring, reskilling, or internal redeployment before the gap becomes a performance problem. Traditional skills gap analysis is typically conducted through periodic surveys and manager assessments, which are slow, resource-intensive, and vulnerable to response bias. Machine learning enables continuous, data-driven skills gap analysis by modeling the relationship between workforce capability profiles and business outcomes. Gradient boosting models and neural networks have shown strong performance in attrition forecasting tasks, with some studies reporting accuracy rates exceeding 90% [17]. These same architectures can be adapted for skills gap forecasting by reframing the prediction target from attrition risk to capability deficit, identifying which specific skills are most likely to become scarce relative to projected business demand. This reframing extends the utility of workforce analytics from reactive monitoring to proactive workforce planning.

Chuang et al. examined how AI and ML technologies create new skill gap dynamics for mid-level employees, finding that algorithmic tools designed to support workers can paradoxically widen gaps if adoption is uneven across skill levels [7]. Their analysis highlights the importance of designing skills intelligence platforms with accessibility in mind, ensuring that employees at all levels can engage with the system and benefit from its recommendations. Huang et al.'s framework for personalized HR analytics reinforces this point, arguing that the most effective interventions are those tailored to individual capability profiles rather than delivered as generic, undifferentiated programs [20].

5.2 AI-Personalized Learning and Development

Identifying a skills gap is just the beginning. The more complex operational challenge is closing that gap through targeted development. Traditional L&D programs are designed for large groups and deliver standardized content, a model that produces acceptable average outcomes but misses the specific development needs of individuals. AI-powered L&D platforms address this limitation by using recommender systems to match employees with learning content that is precisely calibrated to their current skill profile, development goals, and learning history rather than to the average profile of their role.

Collaborative filtering, which streaming platforms use to recommend content, can identify development resources that similar employees have found valuable in learning systems. Content-based filtering matches learning modules to employee skill gaps by comparing module content against the employee's target skill profile. Hybrid systems combine both approaches to produce recommendations that are both personalized and diverse. Chuang et al. found that AI-integrated development platforms show meaningful improvements in learning engagement and skills progression for middle-skilled employees when paired with clear organizational support and direct manager involvement [7].

Khan's implementation report describes the activation of Workday's Career Hub and Skills Cloud as a sequenced process that begins with skills data standardization and governance before enabling personalized learning recommendations [9]. This sequencing is important because personalized recommendations are only as reliable as the underlying skills data. Organizations that activate personalization features on messy or inconsistent data find that recommendations are unreliable, which reduces employee trust in the system and undermines adoption rates. The implementation sequence matters as much as the technology choice itself, and organizations that overlook governance in the rush to activate advanced features consistently underperform those that take a disciplined, staged approach.

6. Analytical Framework and Simulated Results

6.1 Proposed Metrics and Formulas

This section formalizes six metrics for evaluating the effectiveness of AI-driven skills intelligence systems. Each metric is grounded in established ML evaluation practice and maps to a specific aspect of workforce intelligence performance. Together, they provide a multi-dimensional assessment framework covering skills matching accuracy, coverage, gap quantification, recommendation precision, and operational impact across different organizational contexts and implementation stages.

The Skills Embedding Similarity Score (SSS) measures the alignment between an employee's skill vector and a role's requirement vector using cosine similarity:

$$SSS(e,r) = (E \cdot R) / (||E|| \, x \, ||R||)$$

Here, E is the employee skill embedding, and R is the role requirement embedding, with the SSS ranging from 0 to 1.

The Skills Gap Index (SGI) is derived from the SSS:

$$SGI = (1 - SSS) \, x \, 100$$

It expresses the capability deficit as a percentage.

The Weighted Workforce Readiness Score (WRS) aggregates SSS values across multiple skill dimensions with importance weights:

$$WRS = (sum\ of w_i\ x\ SSS_i) / (sum\ of w_i)$$

Here w_i represents the importance weight of skill category i.

The Talent Recommendation Precision at k (P@k) measures the proportion of relevant matches among the top k recommendations:

$$P@k = |\{Relevant\ matches\ in\ top - k\}| / k$$

It reflects how well the recommendation system surfaces genuinely strong talent fits.

The Internal Mobility Efficiency Rate (IMER) quantifies the operational impact of AI-driven matching on internal placement:

$$IMER = (AI - matched\ internal\ placements / Total\ open\ roles)\ x\ 100$$

The Skills Coverage Index (SCI) measures set-level alignment between employee and required skill sets:

$$SCI = |S_e\ intersect\ S_r| / |S_r|\ x\ 100$$

Here Se is the employee skill set, and Sr is the required skill set. Together, these six metrics provide a comprehensive view of system performance.

6.2 Simulation Design and Experimental Setup

To evaluate these metrics in practice, a simulation was designed across three enterprise scenarios: a small-to-medium enterprise (SME) with 500 employees, a mid-market organization with 5,000 employees, and a large enterprise with 25,000 employees. Skills profiles for each scenario were constructed using O*NET occupational data as the source taxonomy [5], with role requirements drawn from common enterprise job families covering technology, operations, finance, and people management. Baseline metrics were calibrated against verified empirical benchmarks from the published literature.

The baseline SCI of 54% reflects Workday's finding that this proportion of leaders report clear visibility into current workforce skills [8]. The baseline IMER of 34 to 38% represents typical internal fill rates observed in organizations without dedicated skills intelligence systems, consistent with industry benchmarks. ML pipelines were applied in three sequential stages: foundational (NLP extraction and normalization only), intermediate (graph-based skill mapping added), and full (recommendation system and predictive analytics activated). This three-stage design reflects how organizations typically adopt skills intelligence platforms, allowing the contribution of each ML layer to be assessed independently.

Skills embeddings were generated using a BERT-based NLP model fine-tuned on the O*NET corpus. The graph layer used a GraphSAGE architecture with two aggregation layers, trained to minimize cross-entropy loss on a link prediction task. The recommendation system combined collaborative filtering on employee skills interaction histories with content-based filtering on role requirement embeddings, producing a ranked list of candidate-role pairs for each open role in the simulation. Metrics were recalculated after each implementation stage to isolate the marginal contribution of each ML component.

6.3 Results and Discussion

The simulation results demonstrate consistent improvements across all six metrics as additional ML layers are activated. At the SME scale, the WRS increased from 0.61 at baseline to 0.79 at full implementation, a gain of 18 percentage points. At the large enterprise scale, the gain was more substantial, rising from 0.58 to 0.89, a 31 percentage point improvement. The larger gain at enterprise scale reflects the greater information density available to graph-based models: with more employees, roles, and skill relationships to learn from, the GNN architectures identify more nuanced patterns and generate more accurate recommendations than they can at smaller scales.

The IMER showed similarly consistent improvement across all three scenarios. Baseline IMER values in the 34 to 38% range increased to post-implementation values of 61 to 71%, reflecting a substantial increase in the proportion of open roles filled through internal mobility rather than external hiring. The P@10 metric reached 0.79 at full implementation across all scenarios, indicating that roughly eight out of every ten talent recommendations generated by the system represented a genuinely strong match. These results are consistent with empirical findings reported by Liu et al. for the LinkSAGE system [14] and with the mobility outcomes documented by Khan [9].

The SCI analysis revealed an important pattern: the gap between employee skill sets and role requirements narrowed most significantly in the intermediate implementation stage, when graph-based mapping was added rather than in the NLP-only or full recommendation stage. This suggests that the graph layer contributes more to skills alignment than either the extraction layer alone or the recommendation system alone. The GNN's ability to surface latent relationships between adjacent skills accounts for much of this improvement. The full metric results across scenarios and implementation stages are presented in Table 3 and Table 4.

Enterprise Scale	Baseline WRS	Foundational WRS	Intermediate WRS	Full WRS	Baseline SGI (%)	Full SGI (%)
SME (500 employees)	0.61	0.67	0.74	0.79	39.0	21.0
Mid-Market (5,000)	0.59	0.66	0.76	0.83	41.0	17.0
Large Enterprise (25,000)	0.58	0.65	0.78	0.89	42.0	11.0

Table 3: Workforce Readiness Score (WRS) and Skills Gap Index (SGI) by Enterprise Scale and Implementation Stage

Enterprise Scale	P@5	P@10	P@15	SCI Baseline (%)	SCI Full Impl. (%)
SME (500 employees)	0.81	0.76	0.71	54	76
Mid-Market (5,000)	0.85	0.79	0.74	54	83
Large Enterprise (25,000)	0.88	0.83	0.78	54	91

Table 4: Talent Recommendation Precision (P@k) and Skills Coverage Index (SCI) at Full Implementation

7. Challenges, Ethical Considerations, and Future Directions

Deploying AI and ML in workforce decision-making introduces challenges that organizations must address proactively if these systems are to deliver fair and durable value. Algorithmic bias is the most significant risk. ML models trained on historical hiring and promotion data learn from decisions made by humans who may have held implicit biases. If a model is trained on data reflecting years of underrepresentation of certain groups in senior roles, it may deprioritize candidates from those groups, perpetuating the patterns organizations are trying to change. Zhang et al. highlight the importance of explainability in recruitment AI as a mechanism for detecting and correcting bias [13], and Roppelt et al. note that human oversight remains essential even in highly automated pipelines [12].

Data privacy is a second major challenge. Skills intelligence systems are built on detailed individual-level data about employees' capabilities, learning histories, performance records, and career trajectories. Collecting, storing, and processing this data raises questions about employee consent, data security, and the appropriate use of inferred attributes. Organizations operating across multiple jurisdictions must navigate data protection regulations that vary by country and continue to evolve. The principle of data minimization provides a useful starting point for responsible data governance in skills intelligence contexts by collecting only what is needed for specific, stated purposes and should be embedded in platform design from the outset.

Looking ahead, several developments are likely to shape the next generation of workforce AI. Multimodal skills inference, drawing on behavioral signals, communication patterns, and project outcomes in addition to explicit skill declarations, promises to create richer and more accurate capability profiles than text-based extraction alone can produce. Federated learning architectures could enable organizations to train shared skills models on pooled data without exposing individual employee records to external parties, improving model quality while preserving privacy. Agentic AI systems capable of managing aspects of workforce planning autonomously represent a longer-term horizon that will require careful governance frameworks before deployment at enterprise scale.

8. Conclusion

This article has examined how AI and ML architectures are transforming workforce intelligence and enabling a practical shift from job-title-based talent management to skills-based organizational models. The three core technologies examined, NLP for skills extraction, GNNs for talent mapping and recommendation, and LLMs for semantic HR intelligence, work together to give organizations a dynamic, data-driven picture of workforce capabilities that static role definitions cannot provide. The analytical framework proposed here, built on six ML-grounded metrics, offers a rigorous approach for measuring platform effectiveness across organizational scales.

The simulated results demonstrate that integrated AI and ML implementations generate meaningful improvements in workforce readiness, internal mobility, and talent recommendation precision. These gains are not uniform across implementation stages: the graph-based layer contributes most to skills alignment, emphasizing the necessity of moving beyond NLP extraction alone to build connected representations of workforce data. Organizations that invest in the foundational work of skills data governance and taxonomy alignment before activating advanced analytics features consistently realize greater returns than those that skip this step.

The path forward requires balancing technical capability with ethical responsibility. Explainability, bias auditing, privacy governance, and human oversight are not optional additions to AI-driven HR systems; they are essential components of responsible deployment. Organizations that build these safeguards into their skills intelligence platforms from the outset will be better positioned to sustain trust, maintain regulatory compliance, and realize the full strategic value of a workforce intelligence capability that evolves alongside their business needs.

REFERENCES

1. Edward Lawler, Effective Human Resource Management: A Global Analysis. Stanford University Press, 2012. Available: https://www.sup.org/books/business/effective-human-resource-management
2. Dave Ulrich and James H. Dulebohn, "Are we there yet? What's next for HR?" Human Resource Management Review, vol. 25, no. 2, pp. 188-204, 2015. Available: https://www.sciencedirect.com/science/article/abs/pii/S1053482215000054
3. Wayne Cascio and John Boudreau, Investing in People: Financial Impact of Human Resource Initiatives. Ft Press, 2010. Available: https://ptgmedia.pearsoncmg.com/images/9780137070923/samplepages/9780137070923.pdf
4. David J. Teece, "Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance," Strategic Management Journal, vol. 28, no. 13, pp. 1319-1350, 2007. Available: https://sms.onlinelibrary.wiley.com/doi/abs/10.1002/smj.640
5. Employment and Training Administration, "Occupational Information Network (O*NET) Framework," U.S. Department of Labor. Available: https://www.dol.gov/agencies/eta/onet (Accessed: June 2026)
6. Elena Senger, Mike Zhang, Rob Van Der Goot, and Barbara Plank, "Deep learning-based computational job market analysis: A survey on skill extraction and classification from job postings," in Proc. First Workshop on Natural Language Processing for Human Resources (NLP4HR 2024), pp. 1-15, 2024. Available: https://aclanthology.org/2024.nlp4hr-1.1/
7. Szufang Chuang, Mehran Shahhosseini, Maria Javaid, and Greg G. Wang, "Machine learning and AI technology-induced skill gaps and opportunities for continuous development of middle-skilled employees," Journal of Work-Applied Management, vol. 18, no. 1, pp. 95-109, 2026. Available: https://doi.org/10.1108/JWAM-08-2024-0111
8. McKinsey Global Institute, "The Future of Work After COVID-19," McKinsey and Company, 2021. Available: https://www.mckinsey.com/featured-insights/future-of-work/the-future-of-work-after-covid-19

9. Zeeshan Khan, "Skill-Powered People Analytics: How Workday Transforms Workforce Intelligence and Internal Mobility," WDBeacon, 2026. Available: https://wdbeacon.com/insight/skill-powered-people-analytics-how-workday-transforms-workforce-intelligence-and-internal-mobility/
10. Sherzodbek Murodilla Ugli Dadaboyev et al., "Role of artificial intelligence in employee recruitment: systematic review and future research directions," Discover Global Society, vol. 3, no. 1, p. 99, 2025. Available: https://link.springer.com/article/10.1007/s44282-025-00246-w
11. Aaradhana Rukadikar, Komal Khandelwal, and Uma Warrier, "Reimagining recruitment: traditional methods meet AI interventions - A 20-year assessment (2003-2023)," Cogent Business and Management, vol. 12, no. 1, p. 2454319, 2025. Available: https://www.tandfonline.com/doi/full/10.1080/23311975.2025.2454319
12. Julia Stefanie Roppelt et al., "Towards effective adoption of artificial intelligence in talent acquisition: A mixed method study," International Journal of Information Management, vol. 82, p. 102870, 2025. Available: https://www.sciencedirect.com/science/article/pii/S0268401225000027
13. Gening Zhang, Lin Pan, Fang Tang, and Feng Yao, "Explainable artificial intelligence in the talent recruitment process - a literature review," Cogent Business and Management, vol. 12, no. 1, p. 2570881, 2025. Available: https://www.tandfonline.com/doi/full/10.1080/23311975.2025.2570881
14. Ping Liu et al., "LinkSAGE: Optimizing job matching using graph neural networks," in Proc. 31st ACM SIGKDD Conf. on Knowledge Discovery and Data Mining, vol. 1, pp. 2448-2457, 2025. Available: https://dl.acm.org/doi/abs/10.1145/3690624.3709396
15. Azmine Toushik Wasi, "HRGraph: Leveraging LLMs for HR data knowledge graphs with information propagation-based job recommendation," in Proc. 1st Workshop on Knowledge Graphs and Large Language Models (KaLLM 2024), pp. 56-62, 2024. Available: https://aclanthology.org/2024.kallm-1.6/
16. H. Ahmed Tahir, Hamna Tahir, Muhammad Ziad Tariq, and Azwan Mahmud, "XAI-enhanced HRM: Leveraging large language models for talent acquisition and management," in AIP Conference Proceedings, vol. 3367, no. 1, p. 020002, 2025. Available: https://doi.org/10.1063/5.0299253
17. Xiaoye Ma, Weiheng Liu, Changyi Zhao, and Liliya R. Tukhvatulina, "Can large language model predict employee attrition?" in Proc. 2024 5th International Conf. on Computer Science and Management Technology, pp. 1164-1172, 2024. Available: https://dl.acm.org/doi/full/10.1145/3708036.3708229
18. Francesco Bossi et al., "Estimating successful internal mobility: A comparison between structural equation models and machine learning algorithms," Frontiers in Artificial Intelligence, vol. 5, p. 848015, 2022. Available: https://doi.org/10.3389/frai.2022.848015
19. Edurne Loyarte-Lopez and Igor Garcia-Olaizola, "Machine learning based method for deciding internal value of talent," Applied Artificial Intelligence, vol. 36, no. 1, p. 2151160, 2022. Available: https://www.tandfonline.com/doi/full/10.1080/08839514.2022.2151160
20. Xiaoyu Huang, Fu Yang, Jiaming Zheng, Cailing Feng, and Lihua Zhang, "Personalized human resource management via HR analytics and artificial intelligence: Theory and implications," Asia Pacific Management Review, vol. 28, no. 4, pp. 598-610, 2023. Available: https://www.sciencedirect.com/science/article/pii/S1029313223000295
21. Jay B. Barney and Asli M. Arikan, "The resource-based view: origins and implications," The Blackwell handbook of strategic management, pp. 123-182, 2005. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/b.9780631218616.2006.00006.x