

Person-job Matching Model Enhanced by Dynamic Job Knowledge Graph

Zongxing Zhang¹, Xiyao Zheng¹, Yixin Yan¹, Jinjia Lu¹, Zhen Hao²

¹*School of Artificial Intelligence and Big Data, Henan University of Technology, Zhengzhou, China*

²*IFLYTEK Co., Ltd., Hefei, China*

Abstract: The persistent two-way mismatch between enterprise recruitment and job seekers' career choices in the labor market, compounded by the fact that technological iteration generally outpaces the renewal cycle of talent cultivation, has placed mounting pressure on precise person-job matching. In the practical operation of recruitment, this contradiction manifests in three key issues: the absence of standardized quality control mechanisms in the fusion of multi-source heterogeneous data, the lack of reliable factual boundary constraints on content generated by large language models, and the difficulty of quantitatively tracking the dynamic evolution of job competency requirements. To address these three issues, this paper proposes a knowledge-graph-grounded large language model enhancement method to alleviate multiple complex problems in person-job matching. The method comprises four components. In terms of data governance, a five-factor quality assessment framework and a multi-field composite similarity algorithm are employed to transform data of various formats into a structured knowledge graph. In terms of temporal evolution, snapshot comparison and trend scoring are utilized to track changes in job competency requirements. In terms of hallucination prevention, the knowledge graph serves as the factual basis for a series of operations—including input filtering, output verification, graph reverse lookup, and credibility scoring—thereby forming a multi-level iterative verification pipeline. In terms of matching, a seven-dimensional weighted scoring system is designed by integrating Holland's vocational interest theory and the Iceberg competency model. Validation experiments were conducted on a standard test set constructed from 100 annotated job descriptions (JDs) and 400 candidate samples, achieving an F1 score of 95.0% for JD parsing, 96.8% for resume

information extraction, and a precision of 92.8% for person-job matching. The experimental results demonstrate that the adopted hybrid entity extraction scheme, which combines domain dictionaries with LLM assistance, achieves an F1 score of 96.8% on the skill extraction task.

Keywords: Knowledge Graph; Multi-Source Heterogeneous Data Fusion; Hallucination Suppression; Dynamic Evolution Analysis

1. Introduction

Zhang Dandan et al. [1], Chen Lin et al. [2] and Yao Jiaquan et al. [3] verified the fundamental judgment concerning the imbalance between talent supply and demand under the digital-economy context from diverse perspectives. Based on case studies of listed manufacturing firms, Wang Yongqin and Dong Wen [4] demonstrated that robots have exerted impacts on China's labor market. Tian Ge and Zhang Xun [5] discussed the influences of digital technologies on agricultural employment from the perspective of the entire national economy. Along with continuous technological evolution, conventional job-analysis and talent-evaluation approaches that have long depended on human-resource practitioners' experience can hardly cope with massive and frequent talent-recruitment demands. The traditional manual screening mode adopted by enterprises suffers from low efficiency, and it struggles to achieve accurate matching for large-scale resumes. By contrast, the intelligent job-talent matching framework constrained by knowledge graph and large-model capability supports automated job-talent matching workflows. It is capable of conducting batch operations including job parsing, skill identification and talent scoring, so as to adapt to scenarios of digital employment services.

Skill requirements vary significantly across different positions, and raw recruitment data

come from diverse and complex sources. Training models with only a single type of data can easily cause data bias. Meanwhile, traditional solutions have multiple limitations: manually sorting out job rules is costly and operates with low efficiency; static knowledge graphs cannot reflect the dynamic changes of industry skill popularity; general large models lack constraints from industry scenarios, and the generated job descriptions are prone to hallucination problems such as fabricated skills. These multiple factors lead to poor practical implementation effects of existing methods in recruitment scenarios. Therefore, this paper proposes a job-talent matching method combining temporal dynamic graphs and domain fact verification. This system aims to fundamentally alleviate the common compound problems in traditional methods, including insufficient matching accuracy, content distortion and missing skill evolution tracking.

Current research on intelligent job-talent matching mainly unfolds along four routes: knowledge graph construction, graph matching recommendation, temporal knowledge evolution, and large model hallucination prevention. For knowledge graph construction in general and talent domains, Liu Qiao et al.[6] and Xu Zenglin et al.[7] built a standardized three-layer graph construction system, and Shen Kejie et al.[8] completed the construction of a static graph in the talent field based on deep learning, yet none of the existing models have introduced the temporal dimension. In terms of graph matching recommendation, Qin Chuan et al.[9] classified existing algorithms into three mainstream approaches: embedding, path and hybrid, and Xu Mengqi et al.[10] introduced static graphs into job-talent recommendation scenarios; however, this method adopts fixed graph weights, making it hard to dynamically adjust matching scores according to real-time changes in skill popularity. As for temporal knowledge modeling, Wang Yuhan et al.[11] and Shen Yinghan et al.[12] conducted research around the general TKG reasoning framework, but their framework mainly targets general knowledge bases and has not been adapted to the actual demand for dynamic tracking of job skills. On the level of LLM hallucination prevention, Tian Yonglin et al.[13], Huang Bo et al.[14] and Liu Zeyuan et al.[15] put forward a complementary framework of general RAG and graph-model integration, while Fang Yanran et al.[16] and Zhao Yixuan et al.[17] sorted out general ideas for AI-enabled

human resources, lacking a dedicated layer-by-layer verification mechanism for recruitment scenarios. The above studies address corresponding segmented problems respectively, but all suffer from the limitations of single application scenarios and the absence of a full-link integrated framework.

Existing graph construction methods are mostly static snapshots and lack standardized quality control and entity alignment processes for multi-source recruitment data. On this basis, this paper designs a multi-source data entity alignment method integrating five-dimensional evaluation and a four-level multi-dimensional filtering pipeline to build a full-link data governance system. Existing hallucination prevention schemes are general frameworks not bound by graph fact constraints in the recruitment field, and this paper develops a four-layer closed-loop verification mechanism comprising input filtering, output verification, graph back-check and confidence evaluation to intercept false information level by level across the whole generation process. Existing temporal models cannot quantify the trends of job skill addition and reduction as well as weight changes, so this paper introduces a hybrid prediction model combining linear regression and moving average to achieve quantitative prediction of the trend of job skill popularity. Existing matching models only take rigid indicators into account and lack the support of vocational psychology theories, and this paper integrates Holland theory and the Iceberg competency model to construct a seven-dimensional weighted matching system. This paper integrates labeled job descriptions and resume samples from multiple industries and builds a standard test set to complete the effect verification of the entire framework.

The contributions of this paper are summarized as follows:

1. A multi-source data entity alignment method integrating five-dimensional quality evaluation is proposed, and a four-level multi-dimensional dirty data filtering pipeline is designed to construct a full-process management and control system covering standardized collection, cleaning and integration of multiple recruitment data sources.
2. A dynamic job-skill knowledge graph based on Neo4j is constructed, and a hybrid prediction model combining linear regression and moving average is introduced to achieve quantitative prediction of job skill popularity trends, assisting

HR in planning talent strategies in advance.

3. With the domain knowledge graph as the factual anchor, structured entities and relations are taken as rigid verification conditions to realize full-process authenticity control of job generation content.

4. A seven-dimensional weighted job-talent matching algorithm integrating Holland vocational interest theory and McClelland Iceberg Model is proposed, combined with a dual-mode AI agent self-learning architecture based on large language models and local knowledge bases.

2. Related Work

The standardized triple-based construction process for general knowledge graph construction is relatively mature[6-7], but such frameworks are more suitable for structured documents and cannot directly process heterogeneous recruitment data in various forms. Although research on talent domain graphs can accomplish resume entity extraction and construct talent portraits[8,18], most of them output static snapshots and lack the temporal dimension, making it impossible to capture the changes in skill demand triggered by industrial iteration. Existing graph recommendation algorithms can be divided into three paradigms: embedding, path and hybrid fusion[9], and some studies have introduced static graphs into job-talent recommendation tasks and achieved favorable results[10]. However, existing solutions generally adopt a fixed weight mechanism, which cannot dynamically adjust matching scores with the annual changes in skill popularity, and has poor adaptability to positions with rapid technological iteration.

Temporal knowledge graphs (TKG) have two major modeling systems, explicit and implicit, with complete representation and reasoning frameworks[11-12]. They are mostly applied to the development of general knowledge bases, but fail to fit the periodic collection characteristics of recruitment businesses. They lack job snapshot storage and quantified indicators for skill demand changes, and thus cannot be directly implemented for the dynamic monitoring of job skills.

To address the issue of large model hallucination suppression, existing studies have analyzed the technological evolution of RAG, the integration framework of knowledge graphs and large models, and the causes of hallucinations[13-15], and some studies have also constructed general

theoretical frameworks for AI-enabled human resource management[16-17]. However, the above prevention measures are all universal solutions, and have not been tailored with a layered graph verification mechanism for scenarios such as JD parsing and skill identification in recruitment businesses. This makes it difficult to resolve large model hallucinations specific to the recruitment field, such as fabricated skills and false salary information.

3. Multi-Source Data Governance and Graph Construction in the Recruitment Field

This paper uses heterogeneous data from multiple sources such as web crawlers, PDF reports, policy documents, social media public opinions and offline datasets as raw data. Through four steps of collection, parsing, cleaning and graph construction, it completes the standardized processing of job data and resume data, and forms a structured job knowledge graph containing nodes and relations.

Multi-source data collection and five-dimensional quality evaluation:

Recruitment data are distributed across five sources, including recruitment website APIs, enterprise official websites, industry PDF documents, policy documents and social media. In terms of page forms, they fall into three categories: static HTML, SPA dynamic pages and document text. Since data from a single source has the problem of one-sided samples, corresponding adaptive crawling methods are adopted: APIs directly parse JSON data; dynamic pages are fully loaded through Playwright; static pages are parsed by simple crawlers. Multi-source data generally contains four types of noise: missing fields, inconsistent name formatting, duplicate collection, and outdated and invalid data. With reference to the big data fusion quality control theory of Meng Xiaofeng et al.[19], a five-dimensional evaluation system covering completeness, timeliness, consistency, anomaly detection and cleanliness is constructed, and the indicator thresholds and weights are shown in Table 1.

Table 1. Five-Dimensional Data Quality Assessment Indicator System

Quality Dimension	Detection Method	Threshold	Weight
Completeness	Skill core field coverage check	$\geq 95\%$	0.25
Anomaly	Anomaly data	No anomaly	0.20

Detection	pattern detection		
Timeliness	Publishing time decay freshness	Valid within 7 days	0.15
Consistency	Field format & enum value check	Full match	0.20
Cleanliness	Data cleaning cleanliness	Standard met	0.20

Four-level multi-dimensional dirty data filtering pipeline: The pipeline is orchestrated via the FilterPipeline class and adopts a pipeline mode to sequentially connect four levels of filters, namely GarbageFilter, SemanticFilter, HistoricalFilter and IndustryFilter, with each subsequent level taking the output of the previous level as its input. The first-level filter removes dirty data by means of regular expressions, blacklists and keywords. The second-level filter identifies repackaged pseudo-new positions based on the similarity calculated by the BERT semantic model. The third-level filter performs temporal stock verification using dual fingerprints of MD5 and SimHash. The fourth-level filter completes industrial scale verification based on skill correlation in Neo4j. Each module supports independent development and testing. During execution, the statistics of passed and rejected data at each level are accumulated. The overall pipeline takes the original job list as input, and outputs the filtered valid job set and the filtering statistics dictionary. The comprehensive data score is calculated according to Formula (1). Data with a comprehensive score above 80 are directly stored in the database and the graph is updated accordingly, while data with a score below 80 are marked for manual review. The cross-source entity alignment rule defines that two job records are identified as the same position if they meet three conditions: the edit distance of company names ≤ 2 , the BERT-based cosine similarity of job descriptions ≥ 0.85 , and the release time difference within ± 3 days. The version with the highest confidence level is selected for storage. The pipeline process is shown in Figure 1.

$$Q(d) = \sum w_i \cdot q_i(d) \quad (1)$$

Hybrid entity extraction method assisted by domain dictionary and LLM: JD and resume texts have problems such as multiple variants of skill terms and ambiguous boundaries. Therefore, this paper proposes a two-stage hybrid entity extraction method based on the "domain dictionary + LLM assistance" paradigm. As shown in Figure 2, in the first stage, a skill synonym set covering six major categories such

as AI and big data is first established. On this basis, regular expressions and keyword matching are used to quickly identify standard skills that frequently appear in documents and conduct the first-round extraction. In the second stage, for long-tail skills that cannot be covered by the dictionary, newly emerging professional terms or entities with ambiguous boundaries, cloud-based LLM APIs are invoked to perform accurate supplementary extraction. In addition, Prompt templates with constraints of job information and entity types are designed to enable large models to output standardized skill names. Finally, the results of the two parts are merged and deduplicated, and synonymous skill nodes such as K8S and JS are uniformly processed using the synonym table. Relation extraction adopts keyword rule matching: proficiency and mastery correspond to REQUIRES, while priority corresponds to BONUS. Skill correlations are constructed through JD co-occurrence frequency statistics, and the graph is imported via batch transactions in Neo4j.

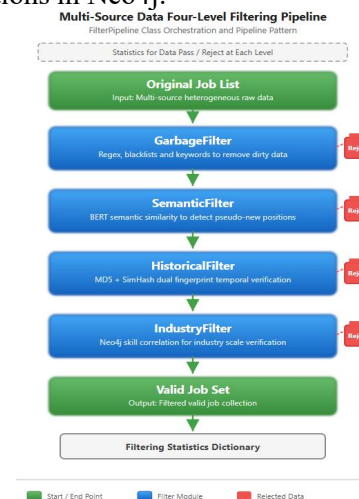


Figure 1. Four-Level Filtering Pipeline for Multi-Source Data

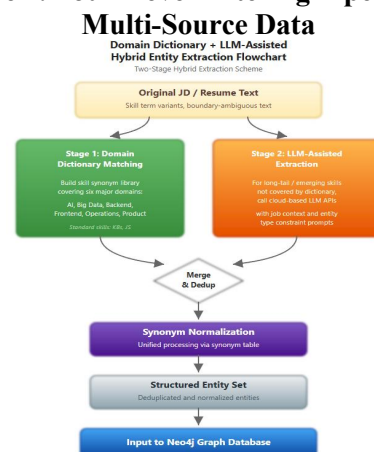


Figure 2. Hybrid Entity Extraction Flow with Domain Dictionary and LLM Assistance

Job knowledge graph construction: The job knowledge graph is constructed using the Neo4j graph database, containing three types of nodes: Job, Skill and Industry, as well as three types of relations: REQUIRES, RELATED and BELONGS_TO. The knowledge graph module automatically extracts job, skill and industry entities and their associated relations, supports batch writing and version snapshots, adopts adaptive weighting of node weights to highlight emerging AI skills, and applies graph database visualization on the front end with drag-and-drop and zoom functions.

Static graph baseline matching scheme: The static graph baseline model only constructs a knowledge snapshot of a single time slice, directly feeds job and skill embedding vectors into matching computation, and completes job-talent scoring solely relying on fixed static association weights. This scheme can only capture current skill associations, cannot adapt to year-on-year fluctuations in skill popularity, and has fixed job matching weights that cannot be dynamically adjusted with industrial technological iteration.

4. Dynamic Graph Enhancement and Intelligent Job-Talent Matching Model

This chapter covers four main parts. Temporal snapshots and trend prediction realize the monitoring of changes in job skill demand and the estimation of their future development. The knowledge graph serves as the factual basis to support the entire four-layer closed-loop verification, covering the whole process from content generation to large model hallucination prevention. Hard skills combined with vocational psychology form a seven-dimensional weighted matching approach. The AI agent integrating LLM and local knowledge bases conducts human-computer interaction after matching and delivers results. Centered on the data flow of the graph, all modules implement a process of "dynamic perception – factual constraint – accurate matching – intelligent interaction".

4.1 Dynamic Skill Knowledge Graph and Temporal Prediction

Previous static graphs and single-layer RAG schemes both fail to demonstrate the year-on-year increase and decrease of job skills and the fluctuation pattern of importance weights. On the basis of the base graph encoder structure, this paper splits out a temporal layered architecture,

dividing the graph into an upper layer of periodic historical snapshots and a lower layer of fixed foundation. The upper layer expands multi-version temporal graphs according to the collection cycle[11][12], while the lower layer maintains the unchanged general job coding structure. The outputs of multi-cycle snapshots are spliced along the temporal dimension to construct a quantitative system of skill evolution, and the overall structure is shown in Figure 3.

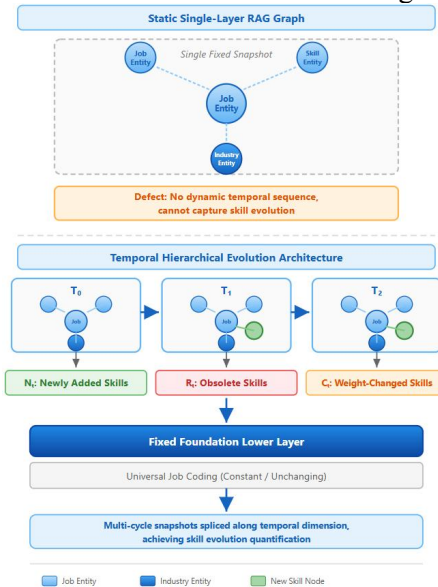


Figure 3. Temporal Snapshot Layered Evolution Model Structure

The model divides the graph into two parts: periodic historical snapshots and a fixed underlying foundation. The first part continuously produces multiple versions of temporal graph snapshots according to the collection cycle, recording the existing skill nodes and the relations between them at each time point. The second part is a constant standard position numbering system used for cross-cycle comparison. The connected results of snapshots from each period present the evolution process of skill importance and their correlation over time, which can be applied to subsequent trend scoring and prediction work.

4.1.1 Temporal snapshot generation mechanism. The system generates complete graph snapshots at a fixed collection cycle, and distinguishes newly added skills, obsolete skills and skills with weight changes by comparing snapshots of adjacent cycles, so as to quantify the intensity of skill changes in a single cycle. The evolution trend scoring is calculated by Formula (2):

$$Trend_score(t) = \frac{|N_t| + |R_t| + \sum_{s \in C_t} |\Delta weight(s)|}{|D_t|} \quad (2)$$

where D_t denotes the total number of positions in

the current graph snapshot, N_t represents the set of newly added skills, R_t represents the set of obsolete skills, C_t represents the set of skills with weight changes, and $|\Delta\text{weight}(s)|$ stands for the absolute value of weight change for skills.

4.1.2 Linear regression trend extrapolation prediction submodule

Based on the temporal data of skill popularity accumulated from multi-cycle snapshots, a linear regression model is adopted to fit the change trend of skill weights, with the formula as follows:

$$\text{weight}_s(t) = a_s \cdot t + b_s \quad (3)$$

where a_s denotes the slope of popularity change for skill s , and b_s denotes the intercept term.

4.2 Four-Layer Closed-Loop Mechanism for Large Model Hallucination Prevention

The basic RAG retrieval-augmented scheme directly injects graph-associated skills into the large model Prompt and conducts single skill keyword matching verification after generation. With only one layer of post-hoc screening mechanism, it is highly likely to fabricate unrecorded skills and made-up salary ranges, and the problem of hallucinations in generated job definitions is prominent.

To limit the degree of generation freedom of large models in recruitment scenarios, this paper adds four layers of verification loops on the basis of basic RAG, namely input boundary constraint, output format and content check, graph entity back-check, and multi-perspective comparison and scoring. Through the above layer-by-layer screening, factually inconsistent content is identified and intercepted. The whole process is shown in Figure 4.

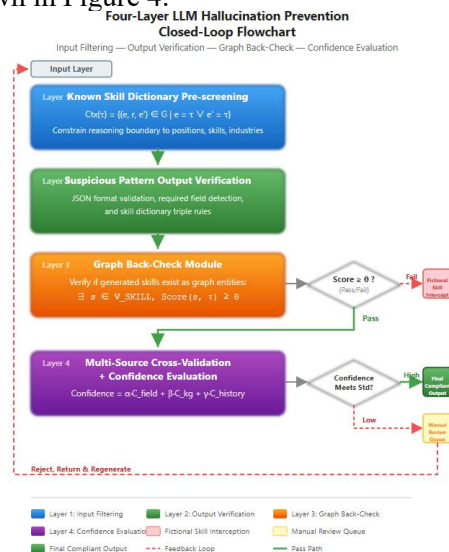


Figure 4. Overall Structure of the Four-Layer Hallucination Prevention Closed Loop

The first layer is the pre-screening module of known skill dictionaries: it retrieves the complete subgraph associated with the target position according to Formula (4), confines the model's reasoning boundary to positions, associated skills and industries, and narrows the scope of generated knowledge from the source.

$$\text{Ctx}(\tau) = (e, r, e') \in G \mid e = \tau \vee e' = \tau \quad (4)$$

The second layer is the suspicious pattern output verification module: it filters abnormal output texts with triple rules of JSON format verification, required field detection and skill dictionary.

The third layer is the graph back-check module: according to the discriminant Formula (5), it conducts reverse retrieval to verify whether the generated skills exist as graph entities, and filters irrelevant fabricated skills through the skill-position matching threshold.

$$\exists s \in V_{\text{SKILL}}, \text{Score}(s, \tau) \geq \theta \quad (5)$$

The fourth layer is the multi-source cross-validation and confidence evaluation module: according to Formula (6), it derives a total score by combining field integrity, graph matching degree and past generation consistency, and content with low scores enters the manual review queue. These four layers together form a complete verification process from input boundary conditions to output fact checking, which largely avoids problems such as false skills and fabricated salaries in the recruitment field.

$$\text{Confidence}(\text{output}) = \alpha \cdot C_{\text{field}} + \beta \cdot C_{\text{kg}} + \gamma \cdot C_{\text{history}} \quad (6)$$

4.3 Seven-Dimensional Weighted Job-Talent Matching Algorithm

Table 2. Descriptions of the Seven Dimensions

Dimension	Weight	Description	Theoretical Basis
skill_breadth	0.25	Skill breadth: coverage of mastered skills	Iceberg Model - Knowledge Layer
skill_depth	0.20	Skill depth: proficiency in core skills	
experience	0.15	Experience: years of work vs job match	Iceberg Model - Experience Layer
project	0.10	Project experience: relevant project match	
education	0.10	Education: degree requirement match	Iceberg Model - Knowledge Layer
culture	0.10	Culture/Values:	

		industry culture match	Self-Concept Layer
industry	0.10	Industry: industry background match	Iceberg Model - Motivation Layer

Combined with the dynamic skill weights output from temporal evolution, a seven-dimensional weighted matching model is constructed, which consists of five hard core dimensions: skill breadth, skill depth, experience matching, project experience and educational background matching, plus two auxiliary dimensions: culture/values and industry matching (corresponding to Holland vocational interest theory[20] and Iceberg competency model[21] respectively). The total score calculation formula is as follows:

$$0.25 \times S_{\text{skill_breadth}} \quad (7)$$

$$0.10 \times S_{\text{project}} \quad (8)$$

The weight allocation, descriptions and theoretical foundations of the seven dimensions are shown in Table 2.

Among them, the skill score is decomposed into two-level indicators of coverage rate and depth matching, realizing refined talent matching scoring under dynamic skill weights, and simultaneously outputting skill gaps and phased learning plans.

4.4 LLM + Local RAG Dual-Mode AI Agent

The AI agent operates in a dual-channel mode combining LLM and local RAG, with functions including job analysis, graph query and trend prediction, and uses WebSocket for streaming conversations to communicate with users. Intent recognition adopts a rule-plus-large-model approach, which can automatically invoke functions such as matching and data analysis, and also has the self-learning capability for high-frequency questions. In terms of retrieval augmentation, the agent can obtain relevant knowledge from the job knowledge graph to assist its reasoning, thereby reducing the possibility of hallucinations generated by large models, and also improving the effects of matching diagnosis and trend prediction.

5. Experiments

The experiments in this paper are conducted on a self-collected multi-dimensional recruitment annotation dataset, covering the following aspects: entity extraction performance evaluation is used to compare the quality of extraction results between the domain dictionary and the

LLM-assisted hybrid method; temporal skill evolution analysis compares the impacts of different snapshot cycles on trend prediction; the seven-dimensional job-talent matching experiment evaluates the matching performance of the weighted model; the overall system performance test summarizes three main evaluation criteria; typical failure case analysis summarizes the causes of matching problems and improvement directions.

5.1 Experimental Environment and Dataset

5.1.1 Experimental software and hardware environment

The whole method is deployed based on Docker containers. The host machine is equipped with an Intel Core i7-13700H processor, 16GB of RAM and a 512GB solid-state drive, and the operating system runs on Windows with Docker container deployment. The data storage module is independently deployed with PostgreSQL 15 structured database, Neo4j 5.x graph database and Redis 7 cache service. Entity extraction and large model reasoning are both completed by invoking cloud-based LLM APIs.

5.1.2 Self-built annotation dataset

This paper constructs a standardized recruitment test dataset, containing 100 manually annotated job descriptions (JDs) and 400 candidate samples, covering six major industry categories such as AI, big data, back-end development and front-end development. All fields of positions, skills, educational background and experience are manually annotated with unified and standardized annotation rules. Text preprocessing rules: Chinese texts are segmented by word segmentation tools to complete field splitting, and a standardized synonym mapping library is constructed for skill terms. Entity annotation uniformly adopts a three-category entity classification annotation[22], distinguishing three types of entities: POST, SKILL and INDUSTRY. The total size of the dataset is shown in Table 3.

Table 3. Recruitment Annotation Dataset Scale

Dataset Type	Sample Count	Core Annotation Content
Test set (positive)	200	Job name, required/bonus skills, education & salary, work years
Test set (negative)	200	Job name, required/bonus skills, education & salary, work years
Total samples	400	JD + candidate paired matching samples

5.2 Entity Extraction Performance Evaluation

This paper adopts a hybrid strategy combining domain dictionary and LLM assistance to conduct skill entity extraction. The specific process is as follows: first, keyword matching is performed using a pre-built domain skill dictionary; for unmatched items, LLM is invoked to perform contextual semantic extraction; then the results are merged and deduplicated to output the skill entity set.

For the skill extraction task of recruitment resumes, this paper evaluates the practical performance of this method based on 400 annotated resume samples, taking the F1-score of the skill field as the evaluation metric. The experimental results are shown in Table 4.

Table 4. Skill Entity Extraction Performance Results

Model/Method	F1-Score
Dictionary + LLM Hybrid	96.8%

The experimental results are consistent with the design objective of this paper. Dictionary matching can effectively cover high-frequency standard skills, while LLM-based extraction can supplement novel skills and synonyms that are not covered by the dictionary. The combination of the two achieves an F1-score of 96.8% on 400 samples, indicating that this combined approach is suitable for skill extraction in recruitment scenarios.

5.3 Temporal Skill Evolution Trend Analysis

Temporal trend analysis is carried out based on the periodic graph snapshot mechanism. Historical snapshot data from four update cycles are selected respectively, and a linear regression model is adopted to conduct trend extrapolation and prediction on the change direction of skill demand popularity.

The experimental results show that more historical snapshots contribute to better trend extrapolation. A single cycle only represents the situation at a single time point and cannot reflect the fluctuations in the intermediate stage, so it fails to well present the real change trend of skill popularity. When the number of snapshot cycles increases to four, linear regression can fit the change slope of a longer time series, thus making more reliable predictions for three scenarios of skill popularity: rising, flat and falling. This also facilitates the identification of newly emerging skills, obsolete skills and skills with weight changes, and further provides guidance for graph updates.

5.4 Seven-Dimensional Job-Talent Matching Experiment

The matching model adopted in this paper adds two auxiliary factors, culture/values matching and industry matching, on the basis of skill breadth, skill depth, experience matching, project experience and educational background matching, forming a seven-dimensional weighted scoring method. Referring to the McClelland Iceberg Model[21] and Holland vocational interest theory[20], this method converts psychological assessment dimensions into quantifiable matching scores, and obtains the total score through weighted summation to judge the degree of job-talent matching.

Matching tests are conducted based on 100 JDs and 400 sets of candidate samples, with the overall matching accuracy as the evaluation metric. The experimental results are shown in Table 5.

Table 5. Job-Talent Matching Experiment Results

Matching Model	Overall Matching Accuracy
Proposed Seven-Dimensional Weighted Matching Model	92.8%

The experimental results show that among the 400 test samples, the proposed seven-dimensional weighted matching method achieves a high matching accuracy of 92.8%. The five hard dimensions strictly verify core competencies and work experience, while the other two auxiliary dimensions further supplement the fit between job seekers and positions, assisting the human resources department in making reasonable job-talent matching judgments in the recruitment process.

5.5 Comprehensive System Performance Test

The comprehensive indicators of three core tasks, namely JD parsing, resume extraction and overall job-talent matching, are counted, as shown in Table 6.

Table 6. System Overall Performance Indicators

Task Type	Indicator	Measured Result	Expected Threshold
JD Entity Parsing	F1-Score	95.0%	$\geq 90\%$
Resume Information Extraction	Skill F1-Score	96.8%	$\geq 90\%$
Full-Process Job-Talent Matching	Matching Accuracy	92.8%	$\geq 90\%$

All three key indicators exceed the expected target of 90%.

5.6 Analysis of Typical Matching Failure Cases

Among the 400 test cases, 29 have matching judgment errors. According to analysis, these errors can be divided into two types: one is for some niche and finely categorized skills, which are less covered in the current knowledge graph, resulting in lower skill coverage and thus underestimation of the capabilities of corresponding talents. The other is for fresh graduates: due to the lack of work experience, the proportion of experience has a large impact on their overall score, while the compensatory effect of the psychological auxiliary dimensions is relatively limited.

Subsequent optimization directions include expanding the skill dictionaries of segmented industries, and adjusting the adaptive weights of the experience dimension for fresh graduate samples.

5.7 Experimental Summary

To sum up, the hybrid extraction method combining domain dictionary and LLM assistance achieves an F1-score of 96.8% in skill extraction, which meets the application requirements. The combination of multi-cycle temporal snapshots and linear regression extrapolation can reflect the trend of skill renewal and change to a certain extent, but the stability of prediction performance is insufficient when the number of samples in a single cycle is small. The seven-dimensional matching scheme proposed based on hard standards and psychological theories achieves a matching rate of 92.8% on 400 test samples, among which the five hard dimensions conduct strict verification on core skills and work experience. The three main performance indicators of the system all meet or exceed the 90% standard: the F1-score of JD parsing reaches 95%, the F1-score of resume skill extraction reaches 96.8%, and the job-talent matching accuracy reaches 92.8%. The AI agent subsystem includes eight modules such as job query, matching diagnosis, graph exploration and data insight. For common problems, it can effectively resolve most issues through approaches including conversation memory and self-learning caching, and can basically cover the business requirements of the entire recruitment process.

The sample size of the existing dataset is limited. In subsequent work, we will expand multi-

industry temporal annotation data, and complete data-driven optimization of matching weights as well as supplementary ablation and comprehensive comparative experiments.

6. Experimental Conclusions

Aiming at three core pain points in recruitment scenarios, this study constructs an integrated job-talent matching framework with graph-constrained large models, and designs four core modules: multi-source data governance, four-layer hallucination prevention, temporal evolution quantification, and seven-dimensional matching scoring.

The complete system implemented in this study corresponds to six core innovations:

1. Multi-source heterogeneous data fusion, which completes data standardization relying on five-dimensional quality assessment and entity alignment.
2. A four-level filtering pipeline that integrates dirty data elimination, disguised fake new position identification, temporal inventory verification and industrial scale verification.
3. A dynamic competency graph based on the Neo4j graph database and temporal snapshot mechanism, supporting periodic tracking of job skills.
4. A seven-dimensional weighted matching algorithm that integrates the Iceberg competency model and Holland vocational interest theory, quantifying psychological dimensions into computable factors.
5. An AI agent-driven self-learning dual-mode architecture covering eight categories of professional capabilities.
6. A new position discovery module adopting a five-dimensional scoring hierarchical classifier to automatically identify emerging positions.

Temporal and matching experiments prove that the more multi-dimensional data from the entire industry chain is integrated into the modeling, the higher the accuracy of position tracking and talent assessment. Compared with static graphs and single-layer RAG methods, the complete framework proposed in this paper achieves a significant improvement in comprehensive performance. In horizontal tests with commercial recruitment analysis tools, it has prominent advantages in entity extraction and matching accuracy under scenarios of segmented technical positions and fresh graduates.

The practical application of this study still faces several practical constraints. The cost of

generating large-scale temporal annotation datasets is high, and the construction of skill graphs for some niche positions is not yet perfect. Meanwhile, the design of matching weights is not entirely based on the algorithm itself and has not completely got rid of the reliance on manual experience. In addition, in-depth ablation studies have not yet been conducted to measure the respective contributions of the two main components: hallucination prevention and temporal prediction. To address the above issues, we plan to make efforts in the following aspects in future work: continue to collect temporal recruitment information from different industries and enrich various skill vocabularies; optimize matching weights using the Analytic Hierarchy Process (AHP); design corresponding ablation experiments to evaluate the effects of the four-layer verification mechanism and temporal snapshots respectively; finally, conduct tests in the actual work of multiple enterprises to examine the generalization ability and adaptability of the model.

References

- [1] Zhang Dandan, Yu Hang, Li Lixing, et al. The Measurement of AI Exposure and Its Impact on Labor Demand in China: Evidence from Large Language Models. *Management World*, 2025, 41(7): 59-75.
- [2] Chen Lin, Gao Yuepeng, Yu Linhui. How Is AI Shaping the Labor Demands of Enterprises? Evidence from Big Data Analysis of Recruitment Platforms. *Management World*, 2024, 40(6): 74-93.
- [3] Yao Jiaquan, Zhang Kunpeng, Guo Lipeng, et al. How Does Artificial Intelligence Improve Firm Productivity? Based on the Perspective of Labor Skill Structure Adjustment. *Management World*, 2024, 40(2): 101-116, 117-122, 133.
- [4] Wang Yongqin, Dong Wen. How the Rise of Robots Has Affected China's Labor Market: Evidence from China's Listed Manufacturing Firms. *Economic Research Journal*, 2020, 55(10): 159-175.
- [5] Tian Ge, Zhang Xun. Digital Economy, Non-agricultural Employment, and Division of Labor. *Management World*, 2022, 38(5): 72-84, 311.
- [6] Liu Qiao, Li Yang, Duan Hong, et al. Knowledge Graph Construction Techniques. *Journal of Computer Research and Development*, 2016, 53(3): 582-600.
- [7] Xu Zenglin, Sheng Yongpan, He Lirong, Wang Yafang. Review on Knowledge Graph Techniques. *Journal of University of Electronic Science and Technology of China*, 2016, 45(4): 589-606.
- [8] Shen Kejie, Huang Huanting, Hua Bolin. Constructing Knowledge Graph with Public Resumes. *Data Analysis and Knowledge Discovery*, 2021, 5(7): 81-90.
- [9] Qin Chuan, Zhu Hengshu, Zhuang Fuzhen, et al. A Survey on Knowledge Graph-Based Recommender Systems. *Scientia Sinica Informationis*, 2020, 50(7): 937-956.
- [10] Xu Mengqi, Xiong Xi, Li Binyong, et al. Construction of Person-Post Matching Recommended System Based on Knowledge Graph. *Application Research of Computers*, 2022, 39(1): 194-198.
- [11] Wang Yuhan, Chen Ziyang, Zhao Xiang, et al. Research Progress and Trend of Temporal Knowledge Graph Representation and Reasoning. *Journal of Software*, 2024, 35(8): 3923-3951.
- [12] Shen Yinghan, Jiang Xuhui, Wang Yuanzhuo, et al. A Survey of Temporal Knowledge Graph Reasoning. *Chinese Journal of Computers*, 2023, 46(6): 1272-1301.
- [13] Tian Yonglin, Wang Yutong, Wang Xingxia, et al. From Retrieval-augmented Generation to SAGE: The State of the Art and Prospects. *Acta Automatica Sinica*, 2025, 51(6): 1145-1169.
- [14] Huang Bo, Wu Shen'ao, Wang Wenguang, et al. KG-LLM-MCom: A Survey on Integration of Knowledge Graph and Large Language Model. *Journal of Wuhan University (Natural Science Edition)*, 2024, 70(4): 397-412.
- [15] Liu Zeyuan, Wang Pengjiang, Song Xiaobin, et al. Survey of Hallucination in Large Language Models. *Journal of Software*, 2025, 36(3): 1152-1185.
- [16] Fang Yanran, Xie Xiaoyun, Shi Junqi. Human Resource Management Research in the Era of Generative Artificial Intelligence: A Work-Flow Analysis Perspective. *Bulletin of National Natural Science Foundation of China*, 2024, 38(5): 820-830.
- [17] Zhao Yixuan, Zhao Shuming, Luan Jiarui. Human Resource Management Based on Artificial Intelligence: A Conceptual Model and Future Research. *Nanjing Journal of Social Sciences*, 2020, (2): 36-43.

- [18] Yang Yanwu, Zhai Panyu, Jin Ying. Entity Recognition and Talent Profiles in Digital Industry Job Based on BERT and BiGRU. *Journal of Wuhan University (Natural Science Edition)*, 2025, 71(5): 667-679.
- [19] Meng Xiaofeng, Du Zhijuan. Research on the Big Data Fusion: Issues and Challenges. *Journal of Computer Research and Development*, 2016, 53(2): 231-246.
- [20] Wu Jianbin, Shen Juanfeng. Analysis of College Students' Vocational Personality Based on Holland's Personality Type Theory. *Education and Vocation*, 2014, (8): 87-89.
- [21] Xu Feng. The Construction of HR Performance Management System: Based on the Competency Model. *Enterprise Economy*, 2012, 31(1): 68-71.
- [22] Zhang Rujia, Dai Lu, Wang Bang, et al. Review of Recent Research Progress in Chinese Named Entity Recognition Based on Deep Learning. *Journal of Chinese Information Processing*, 2022, 36(6): 20-35.