

WHITE PAPER

Science-based matching

Cory Moore, PhD

hirevue[◆]

Organizations invest significant resources in selecting the right people, yet the earliest stage of the talent pipeline receives comparatively little attention. How job seekers discover and connect with opportunities is largely left to passive career sites that offer keyword search and wait for the job seeker to do the rest. Job seekers search for titles and terms they already know, evaluate fit on their own by reading descriptions, and move on if nothing looks right. Their discovery is title-bound, limited to their existing mental model of what jobs exist and what they are called. An experienced warehouse worker searching for "warehouse jobs" may never see a posting titled "Fulfillment Associate," despite being well-qualified for what is fundamentally the same role. The result is a discovery process that selects on terminology rather than capability. Qualified talent visits an employer's career site and leaves without finding the right opportunity. Employers, in turn, are left with noisier applicant pools and missed connections to well-fitting candidates they never knew were there.

Generative AI, grounded in Industrial-Organizational (IO) Psychology principles, makes it possible to fundamentally change this model. Rather than requiring job seekers to search for what they already know, an intelligent system can understand what they can do, build structured profiles of their capabilities and the requirements of available roles, evaluate fit systematically, and surface opportunities they might never have found on their own, connecting talent to opportunity based on capability, not terminology. This paper explores the scientific foundations, methodology, and preliminary validation behind **Science-Based Matching by Hirevue**, demonstrating how it achieves structured, job-relevant, and fair matching at scale.

Overview of Science-Based Matching

Science-Based Matching is an intelligent matching capability within Hirevue's **Match & Apply chatbot**, a talent engagement solution embedded on the employer's career site that engages job seekers early in the funnel, before they apply. Through the chat interface, job seekers share their background, including work experience, skills, career interests, and location preferences, or upload a resume. The system builds a structured profile, compares it against job requirements, and presents the best-fitting opportunities with plain-language explanations of why each job aligns with what the job seeker shared.

The system is powered by a **modular agent architecture**. A matching agent manages the conversational workflow and orchestrates three purpose-built matching tools: **candidate profiling**, **job profiling**, and **fit analysis**. Each tool encodes specific IO Psychology principles and produces structured outputs. The agent coordinates when and how each tool is invoked, ensuring that the matching process follows a defined sequence: gather background, build profile, evaluate fit, present results. This separation of concerns means

each component has a defined scope, can be independently tested, and can be improved without disrupting the others.

The architecture is built on **LangGraph**, a framework for stateful, multi-agent AI systems, and powered by **Anthropic's Claude** foundation large language models (LLM) accessed through **AWS Bedrock**. Claude Sonnet 4.5 handles candidate and job profiling, tasks that require nuanced comprehension of unstructured text and adherence to structured output schemas. Claude Haiku 4.5 handles fit scoring, a more constrained, rubric-based evaluation task that runs in parallel across multiple jobs for efficiency. AWS Bedrock guardrails provide content safety and input/output filtering. All models are pre-trained and are not fine-tuned on customer or job seeker data. Job seeker information is processed only at inference time and is not retained or reused for model training.

Science-Based Matching is **job seeker-facing**, not employer-facing. It does not screen job seekers out, does not provide match information to employers, and does not influence any employer hiring decisions. The job seeker decides which opportunities to explore further and whether to apply. The employer's selection process, including any assessments, interviews, and human review, remains entirely separate and unchanged.

Scientific foundation

Person-job fit

Science-Based Matching is grounded in **person-job fit theory**, the well-established IO Psychology framework for evaluating how well a person's qualifications align with what a job requires. Meta-analytic research has demonstrated that person-job fit is among the strongest predictors of job satisfaction and employee retention, with a meaningful relationship to job performance.^[1] The system operationalizes this framework using the **KSA model** (Knowledge, Skills, and Abilities), a foundational construct in IO Psychology for describing the attributes required for effective job performance. In Science-Based Matching, knowledge, skills, and abilities serve as core competency dimensions, while experience and location function as contextual modifiers that shape how competency fit translates into overall match quality.

The science of structure

A second, equally important scientific foundation underlies the system's design: the **well-established superiority of structured evaluation over unstructured judgment**.

Across decades of research in IO Psychology and decision science, one of the most consistent findings is that structure improves the validity and fairness of human evaluations. When evaluations are guided by defined criteria, consistent rubrics, and

systematic processes, they produce more accurate and more equitable outcomes than when left to unstructured judgment.^[2] This holds not only in personnel selection but across a wide range of prediction and evaluation tasks. Mechanical and algorithmic methods that impose structure on the evaluation process are almost invariably equal to or superior to informal clinical judgment.^[3]

Science-Based Matching applies this principle by **encoding structure directly into the system's architecture**. The matching process is decomposed into modular stages, each with a defined scope, explicit inputs, and structured outputs. Rather than a single model performing opaque end-to-end matching, the process is broken into focused, independently testable steps: profiling the job seeker, profiling the job, and evaluating fit. Each step is constrained to a specific task, and the outputs of one stage become the structured inputs to the next. This is not structure imposed after the fact; it is structure by design.

Methodological rigor in LLM prompt engineering

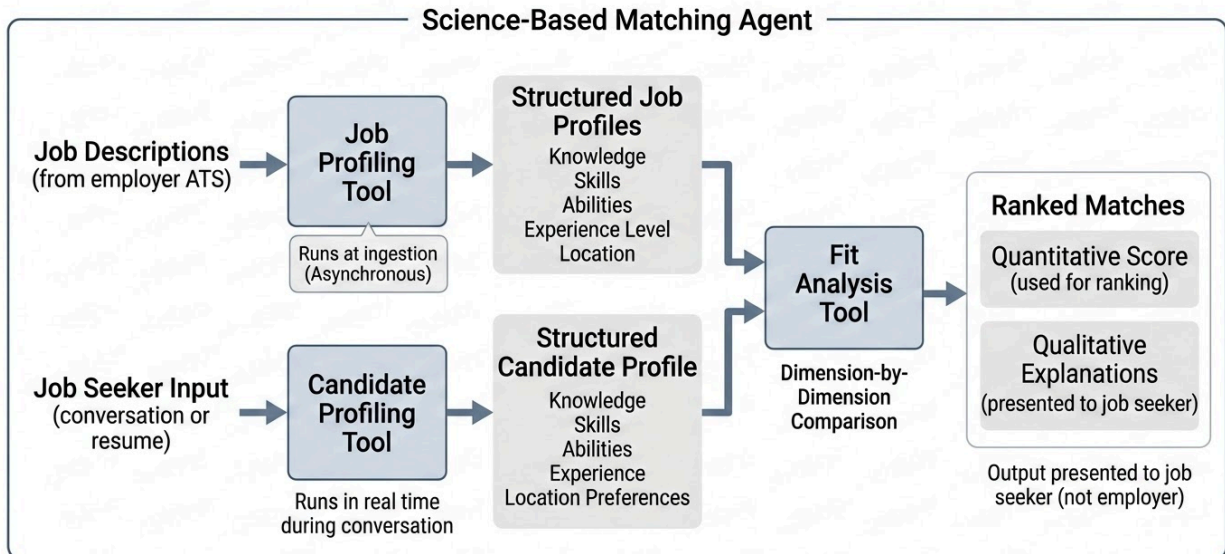
A third scientific foundation addresses how each tool within the structured process is made rigorous. The quality of LLM output is contingent upon the methodological rigor employed in prompt design. Prompts engineered to elicit specific, construct-relevant information produce more valid and reliable outputs than generic, unconstrained prompts. This holds whether the LLM is extracting information from unstructured text or evaluating and scoring against defined criteria.

Science-Based Matching applies this principle systematically across every tool. For profiling, prompts embed structured frameworks (the KSA model), explicit extraction rules, few-shot examples that demonstrate expected outputs, evidence standards for when to infer versus exclude, and defined scope boundaries that constrain the LLM to job-relevant content. For fit analysis, prompts embed **behaviorally anchored rating scales (BARS)** that anchor scores to specific, observable evidence, ^[5] frame-of-reference calibration that defines what each score level means, and explicit evaluation criteria that constrain what the LLM considers and how it weighs it. These are the same psychometric techniques that IO Psychology has shown to improve human rater quality over sixty years of research, applied here to calibrate the LLM.

The convergence is notable. The techniques that improve human evaluators (structured rubrics, defined criteria, anchored scales) and the techniques that improve LLM outputs (rigorous prompt engineering, construct-specific instructions, constrained scope) reflect the same underlying insight: structured tools produce more reliable and valid judgments than unconstrained ones.

Methodology: the matching tools

The matching agent orchestrates three specialized tools, each designed to perform its task the way an expert would: systematically, evidence-grounded, and constrained to job-relevant content.



Candidate profiling

When a job seeker shares their background through conversation or uploads a resume, the **candidate profiling tool** extracts and organizes their qualifications into a structured KSA profile. The tool applies evidence-based extraction rules: it extracts capabilities from direct evidence (e.g., job titles, stated skills), infers capabilities from behavioral evidence (e.g., "led an 8-person engineering team" implies leadership and team management), and excludes capabilities when the evidence is uncertain or insufficient. Extraction follows a two-step process: first, explicit extraction of directly stated qualifications, then systematic inference from accomplishments and responsibilities.

Fairness constraints are embedded at the extraction level. The tool focuses exclusively on job-relevant knowledge, skills, abilities, and experience. It does not extract or consider protected characteristics, demographic information, or demographic proxies. Extracted capabilities are consolidated using canonical, industry-standard terminology to ensure consistent comparison across job seekers who may describe the same skills differently.

Job profiling

The **job profiling tool** applies the same KSA framework to job descriptions, producing a structurally aligned job profile for direct comparison with candidate profiles. This is a form

of **automated job analysis**: the system identifies what knowledge, skills, abilities, experience level, industry context, and work arrangement a role requires, organized into the same dimensional structure used for candidate profiling.

Job profiling runs at ingestion, when jobs are loaded into the system from the employer's applicant tracking system (ATS), not in real time during the job seeker's conversation. Where the ATS provides structured metadata such as experience level or remote work status, those values are used directly rather than being re-derived from the description text, improving data fidelity.

Fit analysis

The **fit analysis tool** performs a dimension-by-dimension comparison of the candidate profile against each job profile. Each dimension is evaluated using **structured evaluation rubrics** with a behaviorally anchored 5-point rating scale capturing the degree of alignment between the job seeker's self-reported qualifications and the job's requirements for that dimension. The rubrics embed expert-level reasoning about what constitutes strong, moderate, and poor fit for each dimension. The rubrics include terminology equivalence instructions so that fit scores reflect actual capability alignment rather than superficial wording differences. For example, "Python programming" and "Python" are recognized as equivalent, and "machine learning" and "deep learning" are recognized as related.

Each dimension is scored independently, and the results are combined using a **compensatory scoring formula**. The overall score begins with the average of the core competency dimensions (Knowledge, Skills, Abilities), which is then adjusted by two weights: **experience-level alignment** (whether the job seeker's career stage aligns with what the role requires) and **location compatibility** (whether the job seeker's location and work-arrangement preferences align with the job's requirements). Each weight scales the competency score based on how well the job seeker's situation matches the job's requirements on that factor. A strong competency fit with a significant experience-level mismatch will be scored lower than the same competency fit with a well-aligned experience level. Similarly, a strong competency and experience fit with an incompatible location will be ranked accordingly. This ensures that practical factors appropriately influence the overall match without being averaged away alongside competency ratings. The formula is principled and transparent, not learned from data.

The fit analysis produces both a quantitative score (used internally for ranking) and qualitative output: key strengths, gaps, and a recommendation presented to the job seeker.

In preliminary validation spanning 200 synthetic candidates and 15 jobs across 5 industry domains, system ratings demonstrated strong convergence with expert-calibrated ground truth ($r = 0.79$, ICC = 0.77 [4]) and strong ranking accuracy (Kendall's $\tau = 0.63$). Areas for continued improvement have been identified, and additional validation with real-world data and larger samples is planned as the system matures.

Bias mitigation and responsible AI

Content safety

AWS Bedrock guardrails provide a critical layer of protection, filtering user input for inappropriate content, managing denied topics (e.g., political content, legal or medical advice), and handling personally identifiable information (PII). The system is designed to **fail closed**: if guardrails are triggered or an error occurs, the system returns a safe fallback response rather than unfiltered content.

Structural bias mitigation

The system's architecture is itself a **bias mitigation strategy**. The profiling tools define a strict extraction scope: they extract only job-relevant knowledge, skills, abilities, and experience, producing structured KSA profiles that exclude demographic information, protected characteristics, and personal attributes unrelated to job qualifications. Fit analysis then operates exclusively on these structured profiles, never on raw resumes or conversation text. The evaluation stage cannot be influenced by information that is irrelevant to job fit, because that information is never present in the data it receives. The bias mitigation is architectural, not instructional — it is a consequence of how the system is designed, not a behavior the system is asked to perform. Preliminary fairness analyses support this design: gender score differences were negligible across all comparisons (Cohen's $d < 0.20$). Score equity across all demographic dimensions, including gender, ethnicity, and age, is assessed through ongoing fairness research, with findings informing continued prompt and rubric refinement.

Transparency and explainability

Science-Based Matching contributes to transparency by providing job seekers with **plain-language explanations** of why each job was recommended. Match explanations describe which skills and experience factors aligned and where gaps exist, grounded in the job seeker's own profile and the job's stated requirements. Job seekers are not shown numerical scores; instead, they receive qualitative descriptions of fit that reference what they shared and how it relates to the role. Because these explanations reference specific

skills and capabilities from the job seeker's extracted profile, they also provide indirect visibility into how the system interpreted the job seeker's background.

The job seeker retains **full decision-making authority**. The system surfaces opportunities and explains its reasoning; the person decides which jobs to explore further and whether to apply. Science-Based Matching is designed to expand access to opportunity, not to limit it. It surfaces jobs that job seekers might not have found through keyword search, such as roles that require their skills but use different job titles or industry terminology.

Summary

Hirevue's Science-Based Matching brings the rigor of IO Psychology to one of the least structured stages of the talent pipeline: how people discover and connect with job opportunities. Built on a modular agent architecture, the system decomposes matching into structured, purpose-built stages for candidate profiling, job profiling, and fit evaluation, each governed by evidence-based rubrics and consistent evaluation criteria. The approach is grounded in two well-established scientific principles: person-job fit theory, which links alignment between a person's capabilities and a role's requirements to performance, satisfaction, and retention; and the demonstrated superiority of structured evaluation over unstructured judgment. Preliminary validation demonstrates meaningful convergence with expert ratings, with ongoing research planned to strengthen the evidence base. Designed for bias mitigation, transparency, and job seeker empowerment, Science-Based Matching complements employer selection processes without replacing any part of them, connecting talent to opportunity where their capabilities genuinely fit.

References

[1]: Kristof-Brown, A. L., Zimmerman, R. D., & Johnson, E. C. (2005). Consequences of individuals' fit at work: A meta-analysis of person-job, person-organization, person-group, and person-supervisor fit. *Personnel Psychology*, 58(2), 281-342.

[2]: Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal of Applied Psychology*, 107(11), 2040-2068.

[3]: Grove, W. M., Zald, D. H., Lebow, B. S., Snitz, B. E., & Nelson, C. (2000). Clinical versus mechanical prediction: A meta-analysis. *Psychological Assessment*, 12(1), 19-30.

[4]: Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290.

[5]: Smith, P. C., & Kendall, L. M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, 47(2), 149-155.

[6]: Zheng, L., Chiang, W.-L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., Lin, Z., Li, Z., Li, D., Xing, E. P., Zhang, H., Gonzalez, J. E., & Stoica, I. (2023). Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena. *Advances in Neural Information Processing Systems*, 36.