ISSN: 2473-4901
v12 n6691

A Multi-Source Analytics Approach to Key Talent Identification: Competency Derivation and Institutionalization in a Digital Firm

Yoonji Kwon
ykwon16@student.gsu.edu

Frank Lee
flee@gsu.edu

J. Mack Robinson College of Business
Georgia State University
Atlanta, Georgia 30303, USA

Abstract

This study introduces a multi-source business intelligence pipeline for objective talent analytics in digital firms, addressing the scarcity of empirically grounded, organization-specific competency criteria derived at scale. Leveraging three distinct data sources—an annual organizational health survey, text-based nomination justifications, and a 360-degree leadership assessment—we apply heteroscedastic t-tests, text mining, and correlation analysis to replace subjective, unscalable nomination workflows with a data-driven approach. The pipeline derives a six-competency framework structured into cognitive, motivational, and social intelligence clusters, with cross-source empirical convergence serving as the primary methodological contribution. Finally, we document the institutional impacts of transitioning to this criteria-based model, including a measurable demographic shift toward earlier-career talent and the subsequent enterprise-wide adoption of the derived framework as the organization's mandatory performance evaluation standard.

Keywords: talent analytics, key talent identification, business intelligence, text mining, competency derivation, digital firms

ISSN: 2473-4901
v12 n6691

A Multi-Source Analytics Approach to Key Talent Identification: Competency Derivation and Institutionalization in a Digital Firm

Yoonji Kwon and Frank Lee

1. INTRODUCTION

This study examines how organizations can identify key talent using internal data rather than managerial intuition. Drawing on three data sources from a digitally driven firm with over 1,000 employees, we empirically derive an organizationally validated competency framework that distinguishes key talent from the broader employee population. Accurately identifying talent is a strategic capability that directly shapes an organization's sustained competitive differentiation (Collings & Mellahi, 2009; Boudreau & Ramstad, 2005). Yet, most organizations lack an empirical basis for this process, relying instead on managerial nomination processes vulnerable to cognitive and structural biases.

Managerial nomination relies on the evaluative judgment of senior leaders. While intuitively appealing, managers holding fixed implicit theories of ability often anchor on early impressions and fail to update assessments (Heslin et al., 2005). Furthermore, managerial potential evaluations are systematically distorted by overconfidence in clinical judgment over actuarial methods (Highhouse, 2008). At the structural level, homophily effects lead managers to disproportionately nominate individuals who resemble them functionally or dispositionally, reflecting managerial similarity rather than organizational need (Golik & Blanco, 2022). The result is a selection process that is neither transparent, consistent, nor reliably predictive.

Competency theory provides a principled foundation for data-driven talent identification, arguing that characteristics predictive of superior performance lie in deep motivational and relational qualities (McClelland, 1973; Spencer & Spencer, 1993; Boyatzis, 2008). However, translating this theory into validated practice remains limited. Empirical studies directly linking specific competencies to performance outcomes using actual organizational data are rare (Boyatzis, 2008). Most applied frameworks rely on expert consensus or generic models without internal validation (Campion et al., 2011), or restrict their scope to specialized executive populations (Kaplan et al., 2012).

This study addresses these gaps. Framed as a business intelligence pipeline, we integrate statistical analytics and text mining to extract objective competency criteria from three data sources: an annual organizational health survey, free-text nomination justification documents, and an external 360-degree leadership assessment. Analytical methods include heteroscedastic t-tests for group mean comparisons, text mining of managerial language, and correlation analysis of survey data. Crucially, the dependent variable—Key Talent versus Non-Key Talent designation—is operationalized using post-calibration outcomes from a cross-functional HR review process rather than initial nominations, thereby removing individual managerial bias from the labeling process.

The study makes two primary contributions: First, it demonstrates that differentiating competency criteria can be derived through convergent multi-source analysis, aligning with the theoretically predicted three-cluster framework of cognitive, motivational, and social intelligence competencies (McClelland, 1973; Spencer & Spencer, 1993; Boyatzis, 2008). Second, it documents the organizational consequences of transitioning to criteria-based identification, including measurable shifts in the demographic distribution of selected talent and the institutional adoption of these competencies as organization-wide performance standards.

The remainder of this paper is organized as follows: Section 2 reviews the literature on strategic talent management, competency theory, and data-driven derivation. Section 3 describes the research context, data sources, and analytical methods. Section 4 presents findings and the integrated competency framework. Section 5 discusses theoretical and practical implications, limitations, and future research directions, and Section 6 concludes.

ISSN: 2473-4901
v12 n6691

2. LITERATURE REVIEW

2.1 Talent Management and the Strategic Importance of Key Talent Identification

The concept of talent management gained widespread organizational prominence following Chambers et al.'s (1998) McKinsey report, which introduced the "war for talent" and argued that attracting, developing, and retaining high-performing individuals constitutes a primary source of sustainable competitive advantage. Since then, scholarship has evolved toward more precise conceptualizations of strategic talent management and its target populations. The foundational definition of talent management as used in this study is summarized in Table A1 in Appendix A.

Collings & Mellahi (2009) offer the foundational definition of strategic talent management, framing it as the systematic identification of key organizational positions that disproportionately drive competitive advantage, the development of a high-potential pool to fill those roles, and the architecture of differentiated HR systems to support them. Central to this framework is the concept of "pivotal talent pools"—the principle that employee contributions are asymmetric, and that concentrated investments in top-tier talent yield superior organizational returns.

Boudreau & Ramstad (2005) extend this resource-allocation logic through "talentship," a decision-science framework for human capital. They argue that achieving meaningful competitive differentiation requires identifying exactly which talent segments are truly pivotal to business strategy.
Taken together, these frameworks demonstrate that effective talent management relies on precisely identifying who within the organization constitutes key talent and establishing the empirical basis on which that determination is made. This study takes that precise question as its point of departure.

2.2 The Limits of Nomination-Based Identification

Despite growing consensus on the strategic importance of identifying key talent, the methods organizations use to make these determinations remain largely subjective. The dominant practice — managerial nomination — has been criticized on multiple grounds. Managers consistently overestimate the predictive validity of their clinical judgment relative to actuarial methods (Highhouse, 2008), and those with fixed implicit theories of ability anchor on early performance impressions, systematically failing to update assessments in response to actual behavioral change (Heslin et al., 2005). At the structural level, managers disproportionately nominate individuals who resemble themselves along functional and dispositional dimensions — a homophily effect that produces talent pools that reflect managerial similarity rather than organizational need (Golik & Blanco, 2022).

This structural problem motivates a criterion-based alternative: rather than relying on managerial intuition, the present study derives selection criteria empirically from organizational data and operationalizes the dependent variable using post-calibration outcomes from cross-functional HR review — entirely removing individual managerial judgment from the labeling process. What is needed, and what competency theory provides, is a structured, theoretically grounded basis for determining who constitutes key talent.

2.3 Competency Theory: What Distinguishes Key Talent

If unstructured nomination-based identification is inherently vulnerable to bias, the question becomes: what observable, theoretically grounded standards can anchor and objectify talent judgment? Competency theory provides the foundational answer.

McClelland (1973) initiated the competency research tradition by challenging the prevailing reliance on intelligence testing in personnel selection. He argued that the characteristics most predictive of occupational success reside not in conventional aptitude measures, but in deeper dispositional qualities—such as motives, traits, and self-concept. Boyatzis (1982) operationalized this theoretical position through large-scale empirical research studying more than 2,000 managers across 41 roles. His findings demonstrated that the competencies distinguishing outstanding from average performers clustered into distinct cognitive, motivational, and interpersonal dimensions. Crucially, technical knowledge proved to be merely a threshold requirement rather than a competitive differentiator. Spencer & Spencer (1993) formalized these insights into the canonical "iceberg model" of competency

ISSN: 2473-4901
v12 n6691

research. This framework distinguishes between surface-level competencies (visible, trainable skills and knowledge) and deep-level competencies (less visible, stable traits and motives).

Revisiting the literature, Boyatzis (2008) confirmed that this three-cluster structure—comprising cognitive, emotional intelligence, and social intelligence competencies—consistently differentiates performance across diverse countries, sectors, and occupational levels. However, he noted an enduring research gap: despite the widespread adoption of competency-based HR systems, empirical studies directly linking specific competencies to performance outcomes using actual organizational data remain scarce. Complementing this lineage, Silzer & Church (2009) approach talent identification from a potential-assessment perspective. They propose a three-dimensional model encompassing foundational dimensions (stable cognitive and personality characteristics), growth dimensions (adaptability and learning motivation), and career dimensions (early role-specific indicators).

2.4 Data-Driven Approaches to Competency Derivation

Traditional competency modeling relies on expert judgment, interviews, and panel consensus—qualitative methods that are resource-intensive, unscalable, and constrained by small sample sizes (Campion et al., 2011). To build robust frameworks, Campion et al. (2011) advocate for convergent validation across multiple independent data channels and anchoring competency derivation in internal performance data rather than normative assumptions. While large-sample empirical assessments can successfully map the underlying structures of performance (Kaplan et al., 2012), a persistent research gap remains. Most organizational frameworks lack direct internal validation (Boyatzis, 2008), focus exclusively on executive echelons (Kaplan et al., 2012), or superimpose pre-existing templates onto single contexts. Though recent work has expanded empirical tracking to broader multinational cohorts (Haro et al., 2023), this study extends the literature by deriving competency criteria inductively across multiple data streams rather than applying a pre-specified framework to a single source.

Parallel Information Systems (IS) scholarship reframes these talent challenges within data-driven decision support architecture. Marler & Boudreau (2017) establish HR analytics as an essential evidence-based practice, noting that talent management remains hindered by manual, judgment-based workflows rather than scalable digital systems. Bottesch et al. (2025) position people analytics firmly within IS research, arguing that the discipline’s dual focus on technology and management makes it uniquely equipped to resolve these persistent talent identification shortcomings.

To date, advanced empirical applications in talent identification remain sparse. While some studies have deployed algorithmic decision trees to model configurations of peer evaluations (Nijs et al., 2022), this paper adopts a complementary approach that balances predictive power with strict methodological transparency and direct interpretability. By simultaneously applying Welch's t-test, TF-IDF text mining, and correlation analysis across three data sources, this study surfaces critical, implicit behavioral signals from naturalistic text that traditional structured surveys miss (Guo et al., 2024; Palshikar et al., 2018). A consolidated summary of these intersecting HR/competency and IS/analytics contributions is detailed in Table B1 in Appendix B.

2.5 Research Gap and Study Positioning

The foregoing review identifies three converging streams of scholarship — strategic talent management, competency theory, and data-driven HR analytics — each of which illuminates a dimension of the problem this study addresses, yet none of which, individually or in combination, has been applied to the specific challenge at hand.

Building on these empirical contributions (McClelland, 1998; Kaplan et al., 2012; Haro et al., 2023), what remains absent from the literature is a study that integrates all three streams: one that applies data analytics methods to derive organization-specific key talent competencies from a full employee population, using convergent evidence across multiple data sources, in a large firm. The present study extends prior work by deriving competency criteria inductively from organizational data across three sources, enabling convergent validation that single-source approaches cannot provide. As Boyatzis (2008) noted, despite 35 years of theoretical development, published studies directly linking specific competencies to performance outcomes within real organizational data remain rare.

ISSN: 2473-4901
v12 n6691

3. RESEARCH DESIGN AND METHODOLOGY

3.1 Research Context

This study was conducted at a digitally driven firm with over 1,000 employees operating at the intersection of technology and digital content. In this environment, competitive differentiation depends on continuous innovation, cross-functional collaboration, and tech-enabled workflows. Previously, key talent was identified through unconstrained departmental nominations lacking standardized evaluative criteria. Selection outcomes varied widely by department, reflecting the implicit standards of individual managers rather than a unified organizational framework. Structurally, this legacy process disproportionately favored senior and managerial employees, creating systematic barriers to identifying high-potential junior talent.

The central research question driving this initiative was: Among employees who receive strong performance appraisals and are nominated by department heads, what common characteristics do they share? Rather than importing a generic competency model, this study inductively discovers the organization's unique talent differentiators from the ground up, using data the firm already holds.
The initiative was led by HR over a six-month period in cross-functional collaboration with key organizational units. The redesigned process reframed talent identification as a population-level binary classification problem where each individual is designated as either Key Talent (Y = 1) or Non-Key Talent (Y = 0).

A critical design choice concerned the dependent variable (Y label): rather than using initial nomination outcomes—which carry severe subjective biases—the study utilized post-calibration outcomes from a cross-functional HR review. This structural choice directly mitigates the anchoring and homophily biases identified by Heslin et al. (2005) and Golik & Blanco (2022) by removing individual managerial judgment from the final labeling process.

The resulting six competency criteria were applied uniformly across the entire employee population, regardless of organizational tier. However, for employees in managerial roles, an additional screening step was incorporated: 360-degree leadership assessment data were analyzed to flag substantive leadership concerns that might undermine a key talent designation. This supplementary step addresses the structurally distinct accountabilities of managerial roles by leveraging an independent, externally administered data source unique to this subpopulation.

3.2 Data Sources

Three distinct datasets were constructed, each addressing a distinct analytical question (Table 1). Their combined use reflects the convergent validation logic advocated by Campion et al. (2011): when findings from multiple data sources converge, the resulting competency framework carries substantially greater evidentiary weight than any single-source analysis could provide.

3.2.1 Dataset A — Organizational Health Survey

The organizational health survey is administered annually to the entire employee population; the survey was fielded in both 2022 and 2023, yielding highly consistent results across both cycles. N = 468 employees completed the survey in the representative cycle analyzed in this study, of whom n = 26 (5.6%) were designated as Key Talent. Items assess organizational commitment, strategic alignment, goal clarity, psychological safety, cross-functional collaboration, and work engagement, using Likert-type scales. The analytical question addressed by this dataset is whether Key Talent employees exhibit systematically different behavioral and attitudinal patterns compared with the broader employee population.

3.2.2 Dataset B — Nomination Justification Text

As part of the organization’s established key talent selection process, department heads have been required to submit a free-text justification for each nominee, covering demonstrated strengths, leadership behaviors, growth potential, and organizational fit. The nomination justifications analyzed in this study were drawn from the entire Key Talent cohort during the most recent completed selection

 ISSN: 2473-4901
v12 n6691

cycle prior to this initiative. These texts constitute a naturalistic record of the implicit criteria managers actually apply when identifying talent — not the criteria they espouse, but those revealed in their language. The analytical question this dataset addresses is: what patterns of implicit evaluative criteria emerge from managerial discourse about talent?

3.2.3 Dataset C — 360-Degree Leadership Assessment

For employees in managerial roles, competency scores were available from an externally administered 360-degree leadership assessment tool that covered multiple leadership dimensions, as rated by supervisors, peers, and direct reports. The analytical question addressed by this dataset is whether selected managerial employees score objectively higher on defined competency dimensions than non-selected peers.

Category	Dataset A: Organizational Health Survey	Dataset B: Nomination Justification Text	Dataset C: 360-Degree Leadership Assessment
Scope	Distributed organization-wide; N = 468 employees completed the survey	All designated Key Talent employees	All managerial-role leaders (excluding newly appointed leaders)
Key Talent Proportion	n = 26 of N = 468 respondents (5.6%)	100% (entire Key Talent cohort)	Key Talent leaders: n = 34; Non-Key Talent leaders: n = 65
Measurement	7-point Likert scale, approximately 30 items	Free-text narrative (authored by department heads)	Multi-rater assessment (self, supervisor, and peer/direct-report ratings)
Timing	2022 (representative cycle)	2022	2022

Table 1. Overview of the three organizational data sources

3.3 Analytical Methods

Because competency theory holds that outstanding performers are distinguished by deep-level cognitive, motivational, and relational competencies rather than surface-level knowledge or perceptions (Spencer & Spencer, 1993; Boyatzis, 2008), a one-tailed Welch's t-test was used to test whether Key Talent employees received higher scores than Non-Key Talent employees on organizational health survey items. This heteroscedastic approach provides unbiased estimates under the unequal group sizes and variances inherent in selective talent data. Of the 30 items included in the organizational health survey, six reached statistical significance; the consistency of these differences across two replication cohorts reduces the likelihood that they reflect chance findings alone.

To complement this quantitative analysis, text mining was deployed on unstructured managerial text to surface latent semantic patterns independent of pre-specified frameworks. Together, these methods mitigate the cognitive anchoring and homophily biases documented in legacy nomination workflows (Highhouse, 2008; Heslin et al., 2005) by deriving criteria directly from empirical data rather than subjective judgment.

A text analytics workflow was executed in RapidMiner on a corpus of approximately 18,000 words of naturalistic nomination justification texts from the most recent selection cycle. Preprocessing included tokenization, stopword removal, and term-frequency weighting via the TF-IDF algorithm to isolate dimension-specific discriminating terms over generic filler language (Guo et al., 2024). Stopword filtering combined a standard English stopword dictionary with a custom organizational stopword list constructed to remove company-specific jargon, internal team and product names, and boilerplate nomination phrasing. Terms were not algorithmically stemmed or lemmatized; instead, morphological variants (e.g., "communication," "communicating," "communicates") were consolidated during keyword review so that the reported terms remain directly interpretable. High-frequency keywords were extracted across three structural dimensions: stated strengths, areas for development, and leadership

ISSN: 2473-4901
v12 n6691

qualities. This design mirrors the large-scale text-segmentation architecture validated by Palshikar et al. (2018) for extracting interpretable performance criteria from open-ended supervisor assessments.

Finally, directional relationships among organizational health survey items were evaluated using correlation analysis. This step served as an internal structural coherence check, verifying whether survey items mapping to the same theoretically predicted competency clusters empirically associated with one another to validate the t-test findings. Supplementary documentation of these methods is provided in the appendices: Table C1 in Appendix C summarizes the strengths and limitations of each analytical approach, and Table D1 in Appendix D situates them within the broader empirical literature on competency derivation.

3.4 Convergent Validation Logic

The three datasets are analytically complementary by design. Dataset A establishes that Key Talent employees are behaviorally distinct from the general population along dimensions consistent with organizational health and engagement. Dataset B reveals what managers actually look for when identifying talent, independent of any imposed framework. Dataset C tests whether the competency differences suggested by the other two sources hold up when assessed with a structured, external, multi-rater instrument. When findings from two or more sources converge on the same competency, the resulting framework achieves convergent validity that no single data source could establish alone. This pattern of cross-source convergence mirrors the best-practice recommendations of Campion et al. (2011).

4. FINDINGS

This section presents findings from each of the three data sources in sequence, followed by their convergent integration into the three-cluster key competency framework. Figure 1 provides an overview of the full analytical process and the convergent validation logic linking the three datasets to the derived framework.

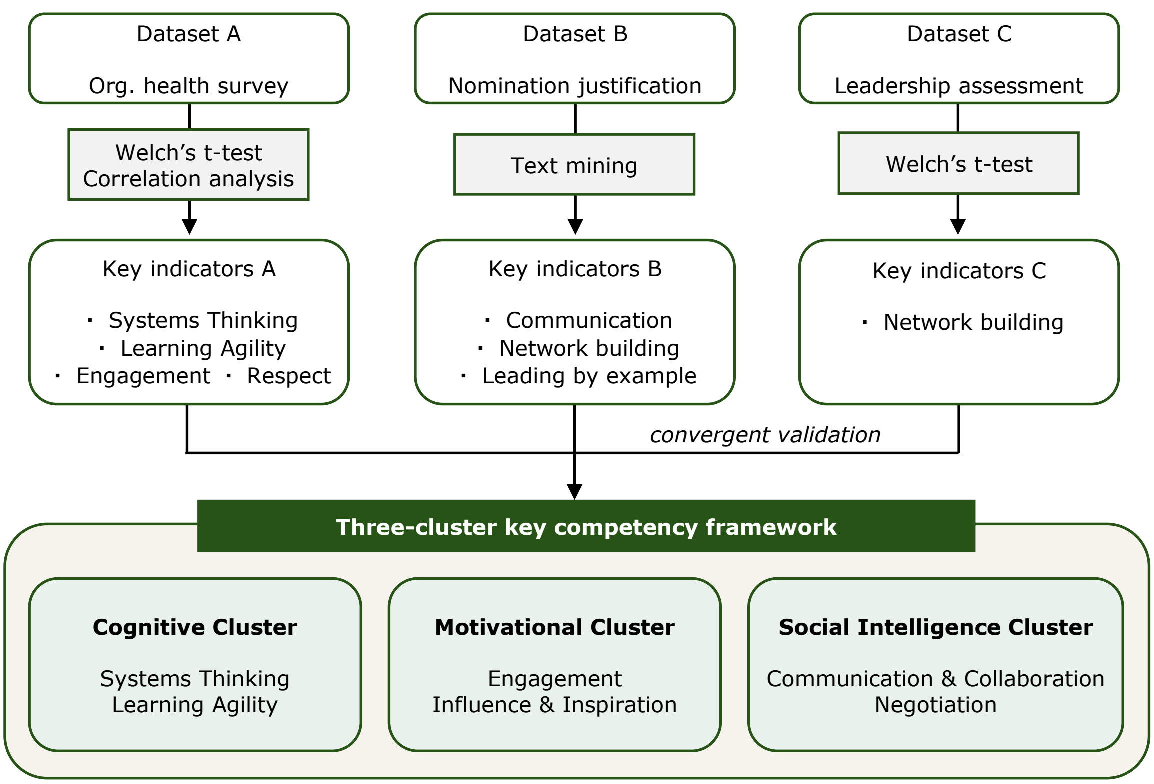

Figure 1. Convergent validation process: three distinct data sources converging on six key competencies.

ISSN: 2473-4901
v12 n6691

4.1 Dataset A: Organizational Health Survey

Competency theory predicts that characteristics distinguishing outstanding performers cluster around cognitive, motivational, and relational dimensions rather than surface-level knowledge or structural perceptions (Spencer & Spencer, 1993; Boyatzis, 2008). To test this, Welch's t-test was applied to an annual organizational health survey across two consecutive cohorts. The same set of items showed stable, statistically significant group differences ($p < .05$) across both years, establishing cross-year replication.

The survey items consistently separating Key Talent from Non-Key Talent clustered into three thematic pillars:
- Strategic Orientation: Vision alignment and mission understanding.
- Motivational Engagement: Goal-oriented achievement, intrinsic task drive, and deep immersion.
- Relational Quality: Mutual professional respect within teams.

Key Talent employees scored significantly higher across all six items, with mean differences ranging from 0.33 to 1.03 scale points. The largest differentiation emerged on vision and strategy alignment ($M = 4.92$ vs. 3.89, $p = .002$). Full item-level statistics are detailed in Table E1 in Appendix E.

Conversely, items measuring perceptions of organizational systems and structural processes failed to reach statistical significance. This concentration of differences within intrinsic motivational and relational attributes—rather than environmental perceptions—directly supports Spencer & Spencer's (1993) theoretical distinction between threshold and differentiating competencies.

Correlation analysis confirmed the structural coherence of this derived model. Survey items mapping to the same theoretically predicted competency clusters were strongly associated empirically. The most robust co-occurrence patterns emerged among items reflecting collective decision-making, openness to perspectives, novel problem-framing, commitment to quality, and principled consistency. These map directly to Communication & Collaboration, Influence & Inspiration, Learning Agility, Engagement, and Systems Thinking, respectively, providing independent structural validation for the overarching three-cluster framework.

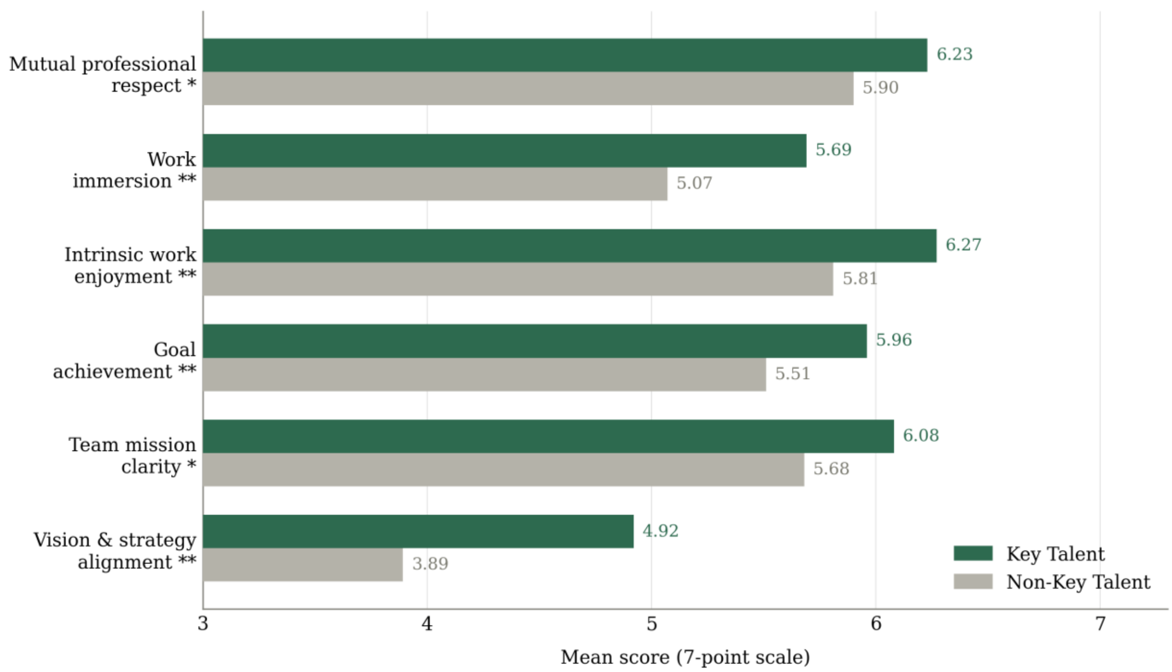

Figure 2. Mean score comparison: Key Talent vs. Non-Key Talent

4.2 Dataset B: Nomination Justification Text — Text Mining Results

ISSN: 2473-4901
v12 n6691

To identify the implicit evaluative criteria managers use when identifying talent, text mining was conducted on the free-text nomination justification documents submitted by department heads during the key talent selection cycle. Documents were analyzed across three dimensions — stated strengths, areas for development, and leadership qualities — and the highest-frequency substantive keywords were extracted for each dimension.

The results reveal consistent thematic patterns across all three dimensions. In the strengths dimension, the most frequent keywords were communication, expertise, and strategy, indicating that managers most readily articulate key talent in terms of communicative competence, domain knowledge, and strategic orientation. In the areas for development dimension, the dominant themes concerned network expansion and communication refinement, mapped to the Negotiation and Communication & Collaboration competencies, respectively. In the leadership dimension, the highest-frequency terms included smooth, facilitative communication, leading by example, and self-sacrifice for the team's benefit — terms that map onto characteristics consistent with the Influence & Inspiration and Communication & Collaboration competencies.

Taken together, the text mining results indicate that managerial language in talent justification documents clusters around two primary themes: strategic and analytical capability on the one hand, and relational and communicative quality on the other. These themes closely align with the cognitive and social intelligence competency clusters in the ultimately derived framework, providing convergent linguistic evidence from a source entirely independent of the survey analysis.

Figure 3a. Keyword frequency from Key Talent nomination justification texts
(Note. Font size reflects relative term frequency across all evaluative dimensions combined.)

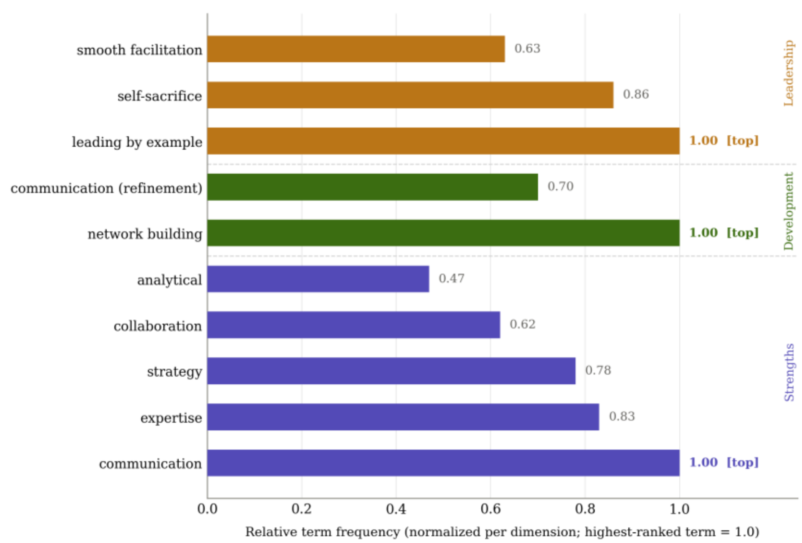

Figure 3b. Relative keyword frequency by evaluative dimension (TF-IDF weighted)

4.3 Dataset C: 360-Degree Leadership Assessment

ISSN: 2473-4901
v12 n6691

For employees in managerial roles, Welch's t-test was used to compare scores from an externally administered 360-degree leadership assessment, contrasting selected Key Talent leaders ($n = 34$) with non-selected peers ($n = 65$) across multiple competency dimensions. Among the dimensions assessed, Network Building — the dimension most directly associated with the Negotiation competency — showed a statistically significant difference between Key Talent and Non-Key Talent leaders ($p < .05$), with Key Talent leaders scoring higher on average than their non-selected counterparts. The remaining dimensions did not reach significance at the $p < .05$ threshold, likely reflecting the relatively small sample of managerial Key Talent employees relative to the comparison group.

Notably, the direction of mean differences was consistent across all assessed dimensions, with Key Talent leaders scoring higher in every case.

4.4 Convergent Validation: The Three-Cluster Key Competency Framework

Drawing on evidence from three distinct data sources, a convergent pattern of differentiation emerges. Key Talent employees are systematically distinguished by their strategic orientation, deep intrinsic motivation, collaborative capacity, and relational leadership. When integrated, these dimensions map onto a coherent three-cluster, six-competency framework.

- The Cognitive Cluster: Captures strategic and analytical differentiators, encompassing Systems Thinking (organically integrated, organizationally grounded decision-making) and Learning Agility (the continuous pursuit of new approaches to generate business value). Evidence derives from survey items measuring strategic alignment and mission understanding, together with correlation-based evidence linking novel problem-framing to Learning Agility, alongside the dominant use of strategic vocabulary in nomination texts.
- The Motivational Cluster: Captures executional and drive-related differentiators, encompassing Engagement (high-intensity commitment, perseverance, and outcome accountability) and Influence & Inspiration (the capacity to inspire peers through exemplary conduct and encouragement). Quantitative survey items measuring Engagement yielded consistently significant group differences, while Influence & Inspiration was independently corroborated through correlation analysis and leadership-dimension language extracted via text mining.
- The Social Intelligence Cluster: Captures relational and interpersonal differentiators, encompassing Communication & Collaboration (building cooperative relationships through respect and attentive interaction) and Negotiation (constructively engaging internal and external stakeholders to achieve shared objectives). This cluster demonstrated the strongest cross-source consistency: communication keywords were the highest-frequency terms across all nomination dimensions, mutual respect items showed significant survey-level differentiation, and network building was the most significantly differentiated dimension within the leadership assessment.

The empirical convergence across three distinct sources constitutes the study's core methodological contribution. This structural alignment supports Boyatzis's (2008) synthesis, which identifies cognitive, emotional, and social intelligence competencies as the enduring differentiators of outstanding organizational performance. Table 2 summarizes the complete integrated framework.

Cluster	Competency	Definition	Primary Evidence
Cognitive	Systems Thinking	Organic, organizationally grounded decision-making integrating business objectives and interdependencies	Org. health survey (strategic alignment, mission understanding); Text mining (strategy keyword)
	Learning Agility	Continuous pursuit of new approaches and knowledge to create business value	Academic literature; Org. health survey (correlation analysis)
Motivational	Engagement	High intensity of commitment, perseverance, and accountability toward achieving outcomes	Org. health survey (performance achievement, work immersion, intrinsic motivation)

 ISSN: 2473-4901
v12 n6691

Cluster	Competency	Definition	Primary Evidence
	Influence & Inspiration	Inspiring peers through leading by example and generous encouragement	Org. health survey (correlation analysis); Text mining (leading by example, sacrifice)
Social Intelligence	Communication & Collaboration	Building cooperative relationships through attentiveness, respect, and considerate interaction	Org. health survey (mutual respect); Text mining (communication dominant across all dimensions)
	Negotiation	Active and constructive engagement with stakeholders to achieve shared objectives	Text mining (network expansion as development area); 360-degree assessment (network building)

Table 2. Key Competency Framework: Three-Cluster Structure and Evidential Basis

5. DISCUSSION

5.1 Theoretical Implications

The findings of this study contribute to three intersecting bodies of scholarship: competency theory, strategic talent management, and data-driven HR analytics.

The derived six-competency, three-cluster framework provides strong empirical support for the theoretical architectures proposed by McClelland (1973), formalized by Spencer & Spencer (1993), and updated by Boyatzis (2008). Specifically, the items that most consistently differentiated Key Talent from Non-Key Talent—strategic orientation, work engagement, and collaborative behavior—align precisely with the deep-level, subsurface characteristics that competency theory predicts will distinguish outstanding performers.

Conversely, surface-level indicators failed to emerge as consistent differentiators across any of the data sources. This reinforces Spencer & Spencer's (1993) assertion that baseline knowledge and skills constitute threshold requirements rather than differentiating competencies. This study thus provides a rare instance of what Boyatzis (2008) noted was missing in the literature: empirical validation linking specific competencies to performance-based group differences within actual organizational data.

This study extends the critique of unstructured nomination-based identification advanced by Highhouse (2008), Heslin et al. (2005), and Golik & Blanco (2022). It demonstrates that a criteria-based alternative is not only theoretically superior but operationally viable. Utilizing post-calibration outcomes as the dependent variable directly addresses the anchoring and homophily biases documented in legacy selection systems. This architecture provides a clear methodological blueprint for organizations seeking to eliminate evaluative subjectivity and improve equity in talent identification.

The convergent multi-source methodology deployed here directly addresses the scarcity of concrete, empirically grounded analytical approaches in talent identification noted by Marler & Boudreau (2017) and Bottesch et al. (2025). Furthermore, it fulfills Campion et al.'s (2011) best-practice recommendation that competency frameworks be derived from multiple data channels rather than constructed through subjective expert consensus. Combining survey-based behavioral data, naturalistic text mining, and external assessment scores offers a robust, replicable analytical template that can be adapted across diverse corporate and industry contexts.

5.2 Practical Implications

Transitioning to criteria-based talent identification generated substantial organizational consequences by expanding the representation of junior and entry-level employees in the selected pool, effectively

ISSN: 2473-4901
v12 n6691

neutralizing the structural seniority advantage of the legacy nomination approach. For corporate succession planning, this earlier identification of high-potential professionals significantly strengthened the depth and demographic diversity of the upstream leadership pipeline.

The impact of this initiative extended far beyond a single talent exercise; the analytics-derived framework functions as the upstream engine of a business intelligence pipeline that cascades directly into six interconnected HR modules to form an enterprise-wide human capital management system. Annually, these uniform criteria combine with performance and leadership data to drive talent identification, feeding a calibrated pipeline that shapes long-term succession planning.

Downstream, the six core competencies were institutionalized as mandatory performance management evaluation criteria across the entire employee population to objectively guide promotion decisions. In compensation and rewards, Key Talent designations trigger targeted salary reviews, explicitly linking discretionary pay and retention to validated behavioral profiles. For talent acquisition, the framework provides structured interviewing protocols that outline specific behavioral prompts and corresponding high-tier response evidence. Finally, learning and development units built customized training pathways around this six-competency architecture, ensuring employee growth remains perfectly aligned with the firm's baseline selection and evaluation standards.

This comprehensive cascade—spanning recruitment, development, performance, compensation, and succession—represents a rare documented instance of an empirically derived people analytics framework achieving full enterprise adoption (Bottesch et al., 2025). Rather than operating as a static research artifact, the competency framework serves as a living, upstream data feed continuously informing decision-making across the full human resources lifecycle. Throughout this ecosystem, strategic HR Business Partner targeting ensures that advisory priorities remain tightly synchronized with the specific organizational functions and individuals where the return on competency investment is maximized.

5.3 Limitations

Several limitations of this study warrant acknowledgment. First, the research was conducted within a single organization in the technology and digital sector, limiting the generalizability of the resulting competency framework. While the three-cluster structure aligns with cross-national competency research, the specific competencies identified may reflect industry- or organization-specific characteristics that would not replicate in different contexts.

Second, while the Key Talent designation incorporated multiple evaluation inputs — including performance appraisal outcomes, leadership assessment scores, and competency criteria — the final determination was made through organizational calibration panels. Although this cross-functional process substantially reduces individual managerial bias relative to unconstrained nomination, panel decisions may still reflect organizational dynamics, such as departmental visibility or positional seniority, that are not fully controlled for in the present design.

Third, the managerial-role analysis was constrained by sample size, limiting the statistical power of the 360-degree assessment comparisons. Future studies with larger managerial populations would be better positioned to detect competency-level differences across a broader range of leadership dimensions.

Fourth, the data are cross-sectional. The nomination justification texts and 360-degree leadership assessment data are drawn from a single 2022 cycle and capture concurrent indicators of engagement, managerial perception, and leadership assessment. The analyses therefore establish that the derived competencies distinguish Key Talent contemporaneously, but they cannot demonstrate the longitudinal predictive validity of the framework—that is, whether these competencies predict sustained leadership and performance success over subsequent five- to ten-year horizons. Longitudinal designs that track identified cohorts through later promotion, performance, and retention outcomes are needed to substantiate the framework’s forward-looking claims.

ISSN: 2473-4901
v12 n6691

6. CONCLUSION

This study addresses a critical gap at the intersection of strategic talent management, competency theory, and people analytics: the absence of data-driven, organization-specific key talent criteria grounded in convergent multi-source evidence across a large employee population. By applying heteroscedastic t-tests, text mining, and within-survey correlation analysis to an annual organizational health survey, text-based nomination justifications, and a 360-degree leadership assessment, this study derives a robust six-competency framework organized into cognitive, motivational, and social intelligence clusters.

The alignment of evidence across three data sources—each capturing distinct dimensions of employee behavior, internal managerial perception, and external assessment—constitutes the primary methodological contribution of this work. Rather than relying on an expert consensus template or a generic literary import, the resulting framework offers a convergently validated characterization of what empirically distinguishes Key Talent within this unique organizational ecosystem.

The study’s secondary contribution lies in its documentation of institutional change. Transitioning from unstructured nominations to criteria-based identification demonstrably shifted the demographic distribution of selected talent toward earlier-career professionals. Furthermore, the derived framework was subsequently institutionalized as the firm's company-wide performance evaluation standard—an outcome reflecting the practical utility and organizational legitimacy that empirically grounded analytics can achieve.

In response to Boyatzis’s (2008) observation that empirical studies directly linking competencies to organizational performance outcomes remain rare, this study offers both a substantive finding and a replicable methodological template. The multi-source analytical architecture employed here is generalizable across corporate contexts. Future research applying similar data-driven designs across different industries, cultural settings, and organizational scales would substantially advance the empirical foundation of competency-based talent management.

7. REFERENCES

Bottesch, S., Schwenke, C., Förster, M., & Klier, M. (2025). Driving business value through people analytics: Literature review and research agenda from an information systems perspective. *Electronic Markets, 35,* 106. https://doi.org/10.1007/s12525-025-00842-3

Boudreau, J. W., & Ramstad, P. M. (2005). Talentship, talent segmentation, and sustainability: A new HR decision science paradigm for a new strategy definition. *Human Resource Management, 44(2),* 129–136. https://doi.org/10.1002/hrm.20054

Boyatzis, R. E. (1982). The competent manager. A model for effective performance. John Wiley & Sons.

Boyatzis, R. E. (2008). Competencies in the 21st century. *Journal of Management Development, 27(1),* 5–12. https://doi.org/10.1108/02621710810840730

Campion, M. A., Fink, A. A., Ruggeberg, B. J., Carr, L., Phillips, G. M., & Odman, R. B. (2011). Doing competencies well: Best practices in competency modeling. *Personnel Psychology, 64(1),* 225–262. https://doi.org/10.1111/j.1744-6570.2010.01207.x

Chambers, E. G., Foulon, M., Handfield-Jones, H., Hankin, S. M., & Michaels, E. G. (1998). The war for talent. *McKinsey Quarterly, 3,* 44–57.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Collings, D. G., & Mellahi, K. (2009). Strategic talent management: A review and research agenda. *Human Resource Management Review, 19(4),* 304–313. https://doi.org/10.1016/j.hrmr.2009.04.001

ISSN: 2473-4901
v12 n6691

Golik, M., & Blanco, M. R. (2022). Homophily: functional bias to the talent identification process? *Personnel Review, 51(2),* 620–643. https://doi.org/10.1108/PR-05-2019-0230

Guo, F., Gallagher, C. M., Sun, T., Tavoosi, S., & Min, H. (2024). Smarter people analytics with organizational text data: Demonstrations using classic and advanced NLP models. *Human Resource Management Journal, 34(1),* 39–54. https://doi.org/10.1111/1748-8583.12426

Haro, J. M. D., García-Izquierdo, M., Castaño, A., & García-Izquierdo, A. L. (2023). Predictive competences of high-potential talent employees. *Journal of Work and Organizational Psychology, 39*(3), 121–130. https://doi.org/10.5093/jwop2023a16

Heslin, P. A., Latham, G. P., & VandeWalle, D. (2005). The effect of implicit person theory on performance appraisals. *Journal of Applied Psychology, 90(5),* 842–856. https://doi.org/10.1037/0021-9010.90.5.842

Highhouse, S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology, 1(3),* 333–342. https://doi.org/10.1111/j.1754-9434.2008.00058.x

Kaplan, S. N., Klebanov, M. M., & Sorensen, M. (2012). Which CEO characteristics and abilities matter? *The Journal of Finance, 67(3),* 973–1007. https://doi.org/10.1111/j.1540-6261.2012.01739.x

Marler, J. H., & Boudreau, J. W. (2017). An evidence-based review of HR analytics. *The International Journal of Human Resource Management, 28(1),* 3–26. https://doi.org/10.1080/09585192.2016.1244699

McClelland, D. C. (1973). Testing for competence rather than for "intelligence". *American Psychologist, 28(1),* 1–14. https://doi.org/10.1037/h0034092

McClelland, D. C. (1998). Identifying competencies with behavioral-event interviews. *Psychological Science, 9(5),* 331–339. https://doi.org/10.1111/1467-9280.00065

Nijs, S., Dries, N., Van Vlasselaer, V., & Sels, L. (2022). Reframing talent identification as a status-organising process: Examining talent hierarchies through data mining. *Human Resource Management Journal, 32(1),* 169–193. https://doi.org/10.1111/1748-8583.12401

Palshikar, G. K., Pawar, S., Chourasia, S., & Ramrakhiyani, N. (2018). Mining supervisor evaluation and peer feedback in performance appraisals. In A. Gelbukh (Ed.), *Computational Linguistics and Intelligent Text Processing* (Lecture Notes in Computer Science, Vol. 10762, pp. 628–641). Springer. https://doi.org/10.1007/978-3-319-77116-8_47

Silzer, R., & Church, A. H. (2009). The pearls and perils of identifying potential. *Industrial and Organizational Psychology, 2(4),* 377–412. https://doi.org/10.1111/j.1754-9434.2009.01163.x

Spencer, L. M., Jr., & Spencer, S. M. (1993). Competence at work: Models for superior performance. John Wiley & Sons.

ISSN: 2473-4901
v12 n6691

APPENDIX A

Table A1. Definition of Talent Management and Competency

Concept	Definition	Reference
Talent Management	The ability to attract, develop, and retain high-performing individuals as a primary source of sustainable competitive advantage	Chambers et al. (1998)
Competency	A capability comprising a set of related behaviors organized around an underlying intent (motive, trait, self-concept); differentiating competencies separate outstanding from average performers	Boyatzis (2008); Spencer & Spencer (1993)

ISSN: 2473-4901
v12 n6691

APPENDIX B

Table B1. Literature Review Summary

This table summarizes key theoretical concepts and empirical findings from the literature review, organized into four major domains: (1) the strategic importance of talent management, (2) the limitations of nomination-based identification methods, (3) competency theory as the theoretical foundation for criteria-based talent identification and (4) IS/analytics approaches to people analytics and talent identification.

Theoretical Domain	Concept	Definition / Core Argument	Key References
Strategic Importance of Talent Management	Talent Management	The ability to attract, develop, and retain high-performing individuals constitutes a primary source of sustainable competitive advantage in knowledge-intensive industries	Chambers et al. (1998)
	Pivotal Talent Pools	Not all employees contribute equally to organizational performance; concentrated investment in those who do yields asymmetric returns. Strategic talent management focuses on identifying who belongs in the pivotal talent pool	Collings & Mellahi (2009)
	Talentship	A decision science framework for HR; meaningful competitive differentiation requires identifying which talent segments are truly pivotal through systematic, data-informed analysis	Boudreau & Ramstad (2005)
Limitations of Nomination-Based Identification	Myth of Expertise	Managers consistently overestimate the predictive validity of their own clinical judgment relative to actuarial, data-based methods	Highhouse (2008)
	Implicit Person Theory (Entity Theory)	Managers holding fixed beliefs about ability anchor on initial impressions and systematically fail to update assessments in response to actual behavioral change	Heslin et al. (2005)
	Homophily Bias	Managers disproportionately nominate individuals who resemble themselves, producing talent pools that reflect managerial similarity rather than organizational need	Golik & Blanco (2022)
Competency Theory: Foundations and Structure	Competency-Based Selection Paradigm	Predictors of occupational success reside not in intelligence scores, but in deeper dispositional qualities (motives, traits, self-concept) that standard tests do not assess	McClelland (1973)
	Competency Definition	A competency is a capability comprising a set of related behaviors organized around an underlying construct called "intent"; differentiating competencies require both observable behavior and inference of underlying motivation	Boyatzis (2008)

ISSN: 2473-4901
v12 n6691

Theoretical Domain	Concept	Definition / Core Argument	Key References
	Threshold vs. Differentiating Competencies	Threshold competencies define baseline adequacy; differentiating competencies (cognitive, emotional, social intelligence) separate outstanding from average performers and are far more predictive of sustained superior performance	Spencer & Spencer (1993); Boyatzis (2008)
	Iceberg Model of Competency	Surface-level competencies (knowledge and skills) are visible and define baseline adequacy; deep-level competencies (self-concept, traits, motives) are less visible and separate outstanding from average performers	Spencer & Spencer (1993)
	Three-Cluster Competency Structure	Competencies differentiating outstanding from average performers converge into three clusters: (1) cognitive competencies, (2) emotional intelligence competencies, and (3) social intelligence competencies	Boyatzis (2008)
	Convergent Validation Best Practice	Competency frameworks should be derived from multiple independent data sources rather than expert consensus alone; convergent validation is the gold standard for competency modeling	Campion et al. (2011)
	Potential-assessment framework	Proposes a three-dimensional model of potential comprising foundational (cognitive/personality), growth (adaptability/learning), and career dimensions; underscores that potential is not unidimensional and effective identification must account for multiple capacity types	Silzer & Church (2009)
	CEO Competency Assessment	Factor analysis of structured assessments of 316 CEO candidates across 30+ competency dimensions reveals two underlying factors linked to firm performance	Kaplan, Klebanov, & Sorensen (2012)
IS/Analytics Stream	Systematic evidence-based review of HR analytics	HR analytics remains in early adoption stages; talent management identified as the domain most in need of systematic analytics application. Most organizations rely on manual, judgment-based processes.	Marler & Boudreau (2017)
	Systematic IS literature review of people analytics	IS discipline uniquely positioned to address people analytics through dual lens on technology and management. Talent identification analytics identified as a priority research gap requiring concrete, well-grounded approaches.	Bottesch et al. (2025)
	Data mining applied to talent identification	Decision trees used to model talent status from peer evaluations. Demonstrates feasibility of algorithmic approaches to	Nijs et al. (2022)

ISSN: 2473-4901
v12 n6691

Theoretical Domain	Concept	Definition / Core Argument	Key References
		talent identification; highlights the potential of data mining methods for deriving talent criteria from behavioral data	
	NLP applied to organizational text data	Demonstrates that classic NLP models (including TF-IDF) and advanced embedding approaches can surface meaningful behavioral signals from HR text data that structured survey instruments alone cannot capture	Guo et al. (2024)
	Logistic regression applied to high-potential competency prediction	Cognitive-intrapersonal competencies (initiative, appetite for learning, thinking beyond boundaries) predict high-potential status in 806 employees. Advances the empirical tradition of competency-based high-potential identification; demonstrates the predictive validity of cognitive-intrapersonal competencies in a large organizational sample.	Haro et al. (2023)

ISSN: 2473-4901
v12 n6691

APPENDIX C

Table C1. Strengths and Weaknesses of Analytical Methods Used for Competency Derivation

Method	Strengths	Weaknesses	Main Application
t-test (Welch's)	Robust to unequal variances and sample sizes; statistically interpretable; replicable across cohorts	Requires numeric data; does not capture content or co-variance among items	Organizational health survey; 360-degree leadership assessment
Text Mining	Captures implicit, naturalistic managerial language; free from response bias; scalable	Frequency-based only; context and sentiment not captured; dependent on preprocessing quality	Key Talent nomination justification texts
Correlation Analysis	Validates structural coherence among competency dimensions; identifies empirical clustering	Does not imply causation; limited to dimensions within the existing instrument	Organizational health survey (competency structure validation)

ISSN: 2473-4901
v12 n6691

APPENDIX D

Table D1. Empirical Methods Used in Prior Competency Studies

Study	Method	Data Source
McClelland (1998)	BEI + group comparison (outstanding vs. typical performers)	Structured interview transcripts
Kaplan, Klebanov & Sorensen (2012)	Factor analysis (30+ competency items) + correlation with firm performance	Structured assessment by external consulting firm (ghSMART)
Haro et al. (2023)	Hierarchical logistic regression + ROC curve analysis	Corporate competency questionnaire + supervisor interview
Present study	Welch’s t-test + text mining + correlation analysis	Organizational health survey, nomination justification texts, 360-degree leadership assessment

ISSN: 2473-4901
v12 n6691

APPENDIX E

Table E1. Organizational Health Survey Items : Key Talent vs. Non-Key Talent Group Comparisons

Item	M (KT)	M (NKT)	SD (KT)	SD (NKT)	df	95% CI	Cohen's d	p	Competency
Cognitive cluster									
Vision & strategy alignment	4.92	3.89	1.65	1.46	27	[0.354, 1.710]	0.66	.002 **	Systems Thinking
Team mission clarity	6.08	5.68	0.98	1.33	31	[−0.018, 0.805]	0.34	.030 *	Systems Thinking
Motivational cluster									
Goal achievement	5.96	5.51	0.72	1.25	35	[0.139, 0.761]	0.44	.003 **	Engagement
Intrinsic work enjoyment	6.27	5.81	0.72	1.31	35	[0.142, 0.772]	0.43	.003 **	Engagement
Work immersion & flow	5.69	5.07	1.12	1.60	31	[0.151, 1.102]	0.45	.006 **	Engagement
Social intelligence cluster									
Mutual professional respect	6.23	5.90	0.86	1.18	31	[−0.038, 0.689]	0.31	.039 *	Communication & Collaboration

Note. KT = Key Talent; NKT = Non-Key Talent (KT: n = 26; NKT: n = 442). One-tailed Welch's *t*-test was used for hypothesis testing; 95% CI reported using the conventional two-tailed critical value for comparability with standard reporting practice. Cohen's *d* computed using the unweighted average of group variances (Cohen, 1988). Scale: 1–7. * p < *.05,* ** p < *.01*.