**American University Kyiv**

A Capstone Project

# AN AI-DRIVEN PREDICTIVE MODEL FOR SCREENING LEGAL PROFESSIONALS IN INTERNATIONALLY ORIENTED UKRAINIAN IT COMPANIES

# ПРОГНОСТИЧНА МОДЕЛЬ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ ДЛЯ ВІДБОРУ ПРАВНИКІВ ДО МІЖНАРОДНО ОРІЄНТОВАНИХ УКРАЇНСЬКИХ ІТ-КОМПАНІЙ

by Volodymyr Dzoban

Presented in Partial Fulfillment of the Requirements
for the Master Degree

APPROVED BY:

Vladyslav Dombrovskyi, Ph.D., Faculty Mentor

2026

**TABLE OF CONTENTS**

## ABSTRACT

This applied research remedies a pressing talents issue for legal departments in Ukrainian IT companies because typical credentials are poor predictors for success in a fast-moving technology sector. The author created and tested a screening tool based on large language model and statistical modelling techniques to better evaluate candidates in an initial assessment stage. Using a dataset of 269 legal professionals who were 178 candidates for a major Ukrainian IT company and 91 professionals in the LinkedIn networking group, this study applied logistic regression analysis and narrative coding to explore predictors for success measured as acceptance of a job offer, retention for 24 months, and delivery of satisfactory performance.

Results demonstrated that a background in the information technology industry, English language skills, foreign transaction experience, interest in technology, and a business focus were strong predictors for success while traditional credentials such as university name, prior employment with a law firm, and membership in a bar association lacked predictive power. The model has 78.4% classification accuracy with 81.2% sensitivity and 75.6% specificity in cross-validation with AUC=0.78. It was used for creating a screening tool based on GPT for candidate screening where output classifies candidate data into organized assessments with scoring points for strong traits, areas of concern, and interview recommendations. Pilot testing showed a 60% reduction in screening time while maintaining quality.

This study addresses a repeatable approach for statistical model development in the context of a particular legal staff environment and a usable screening tool for technology sector legal recruitment. The results of this research challenge traditional forms of credentialism associated with legal recruitment and establish that culture fit factors potentially outperform legal credentials.


*Keywords:* legal recruitment, artificial intelligence, predictive modeling, legal technology, in-house counsel, talent screening, Ukrainian IT sector, GPT, machine learning, human resources analytics

## ACKNOWLEDGEMENTS

I would like to acknowledge that this capstone project would not have been possible without the help and encouragement of several individuals and institutions. Firstly, I would like to extend my warmest gratitude to my faculty mentor for this project, Dr. Vladyslav Dombrovskyi. I am grateful for his thoughtful guidance and leadership inputs in shaping this research project right from conceptualization to completion. His inputs in dealing with the crossover region between technology and other professions have been extremely useful. This has added great quality to this project.

This acknowledgement extends special gratitude to the faculty of the Master of Technology Leadership & AI program at American University Kyiv. Despite the extraordinary difficulties facing Ukraine today, this faculty has created a stimulating environment for learning that is tough as well as visionary. The focus on training technology leaders who can seamlessly merge technical competence with vision is one of the most formative influences shaping this particular research effort.

I would also want to thank the faculty of the Thunderbird School of Global Management at Arizona State University for providing a globally focused curriculum with a strong focus on leadership excellence. This has certainly expanded my horizons with regard to what can be achieved using technology-enabled management.

# LIST OF TABLES AND FIGURES

## Tables



## Figures

## LIST OF ABBREVIATIONS

| | |
|---|---|
| API | Application Programming Interface |
| AUC | Area Under the Curve |
| CIPP | Certified Information Privacy Professional |
| CIPM | Certified Information Privacy Manager |
| CLO | Chief Legal Officer |
| GDPR | General Data Protection Regulation |
| GPT | Generative Pre-trained Transformer |
| IELTS | International English Language Testing System |
| IP | Intellectual Property |
| LLM | Large Language Model (in AI context); Master of Laws (in legal education context) |
| M&A | Mergers and Acquisitions |
| ML | Machine Learning |
| ROC | Receiver Operating Characteristic |
| TOEFL | Test of English as a Foreign Language |

Other abbreviations are used in their standard meanings: AI, AUK (American University Kyiv), Big 4, CV, EU, HR, IT, and UK.

# CHAPTER 1. INTRODUCTION

## 1.1 Background and Context

The IT industry in Ukraine has been identified to be one of the main building blocks of the economy in terms of contributing nearly 5% to the country's GDP during the wartime period (Lviv IT Cluster, 2023). Businesses like Genesis Technologies, MacPaw, and Ajax Systems have developed successful portfolios in the international market to rival the corresponding products originating from leading world economies, without necessarily being located in the USA or Europe. This paradoxical scenario of having a strong international presence while also incurring challenges in the domestic environment poses unique demanding situation for in-house legal counsels of such businesses.

Despite the importance of IT to the country, the tech industry faces difficulties in forming competent legal departments. There is a problem in finding lawyers who have the required profile to be successful in the particular context of tech business. Even if a lawyer comes from a law firm background, that background is not a guarantee of success in a fast-paced tech environment in which the legal professional needs to be concerned with risk management of the entire business. The consequences of mistakes in forming the legal department go beyond the related operational cost and may include effects on the country's reliability in the tech industry as a whole.

The last two decades have witnessed a marked shift in literature trends regarding personnel selection. Schmidt & Hunter (1998) introduced the standards for criterion-related validity for methods of selection. Nevertheless, Sackett et al. (2022) later modified such standards and made a claim that conventional methods were less accurate for predicting success in a role compared to previously thought. Simultaneously, a paradigm shift towards evidence-focused talent management best practices using data analytics instead of management insight has taken place. This shift has been discussed and recommended in a literature study regarding contemporary trends in talent management issues by Cappelli & Keller (2014). These trends taken together show a point in time when new selection science and ML can be applied to the problem of legal hiring in the tech industry.

The effort of Ukraine to integrate into the EU makes the challenge faced by the country even more urgent. With the country trying to synchronize the legal systems to harmonize with the EU rules, having well-staffed legal teams become even more vital in terms of compliance for different jurisdictions. Legal experts in the tech industry in Ukraine need to show expertise in the country's laws in general but also in the topics of GDPR, international arbitration, cross-border data transfers, and sector-specific regulations affecting financial technology, advertising, and consumer protection. The question of how to identify candidates capable of mastering this complexity during the hiring process has become a strategic priority.

## 1.2 Problem Statement: The Hidden Cost of Inefficiencies in Legal Hiring

In Ukrainian IT companies, the legal hiring faces the challenge of screening and interviewing being inherently ineffective. In the traditional screening and interviewing process of candidates, the technique of screening resumes for the evaluation of prospective employees falls short in distinguishing between the candidates who would succeed in a legal position and the ones who would fail. This inefficiency takes two major forms. First, there is a high attrition rate in the new hire pool since a high number of employed individuals fail or leave the firm soon after being hired. Second, the screening process itself consumes substantial management time that could otherwise be allocated to strategic legal work.

The scope of the problem can be appreciated by examining the numbers in the context of real-world recruiting. Citing my personal experience in the position of CLO at TechCo[1], there are about 1.4 interviews (including bar-raising interviews) every week averaging 134 minutes per session. This equates to about 73 interviews per year and results in successful new talent induction in the business in the form of 15 new members annually, resulting in a conversion rate of 20.5%. The challenge here clearly underlines the likelihood of perhaps up to 80% of the process being the 'dark pool,' that is, activity in the form of interviews being unproductive to the degree of about 129 hours (3.2 full-time working weeks) per year.

The inefficiency also imposes costs that extend to more than the mere loss of time. There are costs that directly relate to the failure to hire the required personnel. They include recruitment fees, onboarding investments,

---

[1] TechCo is a pseudonym for a Ukrainian IT company where the author serves as Chief Legal Officer. Pseudonymization is used to protect proprietary hiring data against automatic search engine indexing and data crawling.

and severance payments (if any). Other, indirect costs would be the opportunity cost of the high-level lawyers spending their entire time doing interviews instead of practicing law. There would also be the loss of efficiency in the teams operating below their full capacity for a considerably long period, and the institutional knowledge deficit created by turnover. Studies show that employee turnover has a negative effect on the performance of an organization, and this effect is realized not only through replacement costs but also through the disruption of processes and loss of an organization's cumulative knowledge (Ton & Huckman, 2008). Finally, when a legal hire does not work out, the entire hiring process must restart, compounding these costs.

This research aims to solve the problem of the lack of evidence-based criteria for candidate screening in the context of legal staffing. The current approach to candidate selection relies largely on personal analysis of resumes and unstructured interviews, even though extensive research has shown that these methods have relatively low predictive validity compared with alternative selection tools (Schmidt & Hunter, 1998; Sackett et al., 2022). Additionally, empirical studies about hiring discretion have shown that subjective decisions made during screening sessions affect recruiting results and that such discretion can cause discrepancies in making hiring decisions (Hoffman et al., 2018).

The existing gap between the principles of selection science and the reality of the hiring process provides a challenge for improvement by incorporating ML algorithms to identify patterns in successful legal candidate selection that may not be effectively considered by human assessors.

### 1.3 Research Objectives and Questions

The present capstone project explores the main research question: "Can ML analysis of successful legal hire profiles in top Ukrainian IT companies establish an efficient predictive model for filtering law candidates in order to decrease the screening-to-hire time and improve the accuracy of selection?" The research question connects the theoretical views in the field of personnel psychology to the application of ML to tackle a specific management issue.

Objectives of the research can be distinguished in three different dimensions.

The empirical one is the need to build the first systemic dataset on the characteristics of the legal talent pool and their results in the context of the Ukrainian IT industry. This dataset would include both quantitative information (educational background, years of service, language skills, ex-employer type, etc.) and a qualitative one (traits of a professional path, indicators of motivation, indicators of a good fit, etc.) provided in the resumes.

The second dimension involves methodology itself: the design of a screening tool incorporating AI specifically targeted at the legal roles in tech companies. This comprises the search for the parts of the CV required for a prediction of success, the analysis of different techniques (logistic regression analysis, decision trees, or classification techniques in the form of a large LLM classification model), and defining the scope of validity of the tool over different cross-validation steps.

The third objective is practical in nature – develop a tool for legal managers of the Ukrainian IT industry, to be made operational in a form of improved candidate selection. The target output of the development process would be a tailored candidate screening tool based on GPT to analyze the CV inputs in order to produce a level of suitability that allows selection managers to focus more on interviewing the shortlisted candidates instead of the rest.

Secondary research questions form the basis for the main one. They include the following: Which of the factors in a CV best predict a positive outcome of the legal hiring process – ranking of the university attended, prior employment in IT companies, level of English proficiency skills, or career path type? To what degree may the quality factors of a CV add to the predictive capabilities of a candidate's quantitative skills? Is there the possibility of developing an ML model to be applicable in multiple IT companies in Ukraine? Additionally, how should the application of automated hiring systems in Ukraine be constrained from a legal and ethical standpoint?

### 1.4 Significance of the Study

Similarly to the objectives, this research makes a three-fold contribution: theoretical, methodological, and practical. The theoretical aspect of research expands existing theories of Human Capital Theory (Becker, 1993) in combination with Signaling Theory (Spence, 1973) and Person-Environment Fit Theory (Kristof-Brown et al., 2005) to a new context of legal recruitment in the emerging market tech companies. Those theories have a robust validation in the developed markets but have the potential to be a point of reference in Ukraine's IT business as well.

Methodologically speaking, the research contributes to the growing literature on the application of AI in the management of human resources. Based on the conceptual frameworks developed by Brynjolfsson & Mitchell (2017) to search for tasks to automate by ML algorithms, the research applies these frameworks in the context of screening legal candidates. The research methodology combining both quantitative logistic regression analysis and qualitative narrative analysis as well as classification via an LLM provides a template for further research to quantify the processes of making hiring decisions.

At the practical level, the research satisfies the current needs of the IT industry in Ukraine. A total of 129 hours of unproductive interviewing time estimated previously corresponds to the case of a single legal manager. Taken together for the industry as a whole, the efficiency gains from the improved initial screening may be significant. More importantly, the quality of the legal teams provided by improved hiring decisions would be higher to allow for successful business development in the context of the competitive world market.

The relevance of this study goes beyond the legal function. Indeed, the approach of the extraction of predictive patterns from the past selection of candidates followed by the application of AI tools could be applied to other highly specialized jobs in the tech industry. These jobs include product managers, data scientists, compliance officers, and many more whose selection may not rely on the conventional approach. Therefore, this research aims to contribute to the wider conversation about how organizations can apply AI technology to improve talent decisions.

There is also a connection to the strategic positioning of Ukraine. Given the already mentioned country's ambition to integrate into the EU and become a trustworthy partner in technology, the quality of the in-house legal capabilities of a given business would appear to be a competitive edge. Such organizations would be in a more favorable position to deal effectively with complex regulation, obtain advantageous terms in their contracts, and counteract the negative effects of compliance failures that hurt the image of the business. Thus, the current research aligns with the overall objectives of economic development of Ukraine.

### 1.5 Scope and Structure of the Capstone

The capstone involves the legal hiring process in internationally focused IT companies in Ukraine that develop products for foreign clients. This includes dealing with multijurisdictional laws. The research does not involve the hiring process of the legal staff in domestically focused companies, law firms, or the legal departments of the government. This emphasis is the result of my professional interest and the strategic importance of the export IT sector to Ukraine's economy.

The primary data source is a sample of legal professionals' profiles in hiring of which into TechCo I have participated personally. For them complete hiring records and performance outcomes are available. This sample is supplemented by data of those who were considered to TechCo but were not hired and by data from the LinkedIn professional group established by me, which includes professionals interested in in-house legal practice at Ukrainian IT companies[2]. Together, these sources provide sufficient data for pattern identification while reflecting the actual population of interest. If methodologically necessary, the sample may be further extended to include more legal hires in future studies.

The scope of the research includes the development of specific techniques and a screening tool while excluding a number of related subjects. For instance, the research does not relate to the optimization of the technique for carrying out interviews, the design of the program for new employees, or the organizational form of the legal department. These all affect the selection of new staff but fall outside the scope of this particular study. Additionally, even though the research discusses the legal and ethical implications of the use of AI in the hiring of new staff members, the research does not offer the legal opinion on compliance with the stated jurisdictional demands.

[2] "Disruptive Lawyers Ukraine" LinkedIn group accessible at https://www.linkedin.com/groups/14800108/.

# CHAPTER 2. LITERATURE REVIEW

## 2.1 Theoretical Foundations

A predictive hiring model requires a strong theoretical base that explains the rationale behind the connection between particular characteristics of a candidate and their job performance. Three different theories taken together form the theoretical base for the predictive hiring model proposed in this work. These theories include the Human Capital Theory, the Signaling Theory, and the Person-Environment Fit Theory.

### *2.1.1 Human Capital Theory (Becker)*

As outlined in Human Capital Theory, conceptualized by Gary Becker in 1964 (expanded in 1975 and 1993), a grasp of foundational economic concepts regarding why education and work experience remain critical in the selection process has been identified. According to this theory, individuals make investments in their human capital based on a forecasted increase in future job opportunities, with the output of such investments generating productivity and additional income. Human capital can never be separable from its owner.

In the current research, the Human Capital Theory provides the justification for the analysis of educational background and work experience to determine their predictive capabilities for success in the selection of lawyers for TechCo. When the model determines how lawyers who graduate from top-tier educational institutions or have a particular type of background in terms of their work experience perform well at TechCo, the model indirectly validates the theoretical predictions made by Becker on the topic of human capital. According to the theoretical construct, factors such as the quality of the educational institution attended, the capability in the English language, or the background in the IT industry should act as very strong predictors of success.

The theory relies on some assumptions that must be acknowledged. Firstly, individuals make rational decisions regarding their education based on anticipated career benefits. This may not always be the case. Secondly, this theory assumes that enhanced productivity brought about by learning achieved via education and work experience benefits individuals because their employers have the capability to measure this. While such assumptions are valid in a developed economy, generalization with regard to the Ukrainian legal environment in the IT field would be explored in this empirical stage of this research.

### *2.1.2 Signaling Theory (Spence)*

Unlike the viewpoint of Human Capital Theory, which focuses on the productive quality of education and skill, Signaling Theory, conceptualized by Michael Spence in 1973, offers an alternative explanation for employers' reliance on resumes and educational credentials. Signaling Theory explains how education and resume information serve as signals that allow employers to identify high-ability workers when they cannot directly observe a person's productivity before hiring. Michael Spence was awarded the Nobel Prize in Economics in 2001 (along with George Akerlof and Joseph Stiglitz) for his groundbreaking work on information asymmetry in markets.

This theoretical framework has very clear implications for the current research. In assessing the variables of university rank, past employers, or certification level, the framework effectively tests how well a particular combination of information best correlates to success in the legal profession in the IT industry. Spence's theoretical framework explains how particular facets of a CV will sometimes correlate to the activity despite a lack of direct application to the subject. For example, a degree from a top law program may reflect problem-solving skills, work ethic, and drive even if the particular courses taken have a low direct application to IT law.

The theory also explains employer behavior that could seem unreasonable when viewed from other perspectives. In screening potential employees based on university prestige or corporate affiliation, screening managers use such factors as heuristics to signal talent because of information asymmetry. While using heuristics as a means for drawing inferences about potential employees may seem unreasonable when considered based on other theories because heuristics are not perfect methods of inference, using heuristics can be reasonable because the signal-to-noise ratio may be positive. However, Spence's model also predicts that a positive signal may become less informative with more applicants holding similar credentials. This means that better-performing individuals often seek more costly signals to be distinguished among similar applicants.

The implications of the assumptions in the Signaling Theory have to be reviewed as well. The Signaling Theory states that the cost of gaining credentials for high-ability professionals would be less compared to others –

a possible but not absolute hypothesis. Employers would also be aware of the above reality and factor the costs of the signals into their staffing decisions. In scenarios where the above hypotheses may not hold true, signals may carry less information than the Signaling Theory indicates. The empirical analysis in the research will determine the signals that carry information in the context of the IT legal environment in Ukraine.

#### *2.1.3 Person-Environment Fit Theory (Kristof-Brown)*

Though Human Capital Theory and Signaling Theory consider candidate characteristics, they fail to elucidate why some well-credentialed candidates fail while others succeed despite less distinguished backgrounds. Person-Environment Fit Theory addressed this limitation by focusing on the match between individuals and their corresponding work environments. The theoretical construct was initially integrated into a conceptual framework in Kristof's (1996) pioneering literature review, and later comprehensively examined in the meta-analysis by Kristof-Brown, Zimmerman, & Johnson (2005).

One of the major insights that can be taken away from this literature is that subjective fit, which refers to the level of perceived fit between applicants and organizations, often has a more powerful effect than more objective measures of fit. Perceived fit between applicants and a potential employer influences applicants' initial attitudes towards the organization and subsequent job-related behavior. This has very important implications for recruitment because how applicants advertise themselves in their CV and how applicants state their career goals influences perceived fit with regard to the organization.

In the present research, the role of Person-Environment Fit Theory is a guiding framework that attempts to conceptualize the reasons behind the success of some lawyers in TechCo compared to others having a similar background. When the framework analyzes professional goals, indicators of culture, themes of motivation in the context of CV analysis, the main point of the model attempts to predict person-organization fit. At the same time, the role of Person-Environment Fit Theory suggests that the success of the hiring process should not be determined by the professional background of the candidate only but also by the expectations of the candidate to succeed in the particular context of a high-tech corporation in Ukraine.

The theory assumes the existence of values and preferences that remain constant for the person being considered for a position. These values and preferences can then be compared to the traits of the business. Additionally, the ideal fit can then be determined to a reasonable degree of accuracy in the hiring process. These conditions have been assumed to have some validity in empirical studies. In the fast-evolving tech industry, the determination of fit may be made difficult given the change in the business environment to which the candidate pledges allegiance in the present compared to the one they will face in the future.

### 2.2 Personnel Selection Methods: Validity and Utility

The theoretical frameworks mentioned above have explained the reasons for anticipating the predictive accuracy of particular characteristics of a candidate for a particular type of job; but they neither estimate the accuracy of different prediction techniques nor quantify their potency. These aspects have been extensively researched in the field of personnel psychology by way of some landmark studies known as meta-analysis.

Schmidt and Hunter (1998) made the definitive analysis of the last eighty-five years of research in personnel selection. Their meta-analysis incorporated the conclusions of hundreds of studies to determine what factors have been shown to have predictive validity in selecting for future performance on the job. Their dataset established a clear order of magnitude of validity: the highest was work sample tests at a predicted validity of 0.54, followed by cognitive abilities tests (0.51) and structured interviews (0.51), integrity tests (0.41), unstructured interviews (0.38), and years of employment in a given position (0.18).

Recently, a very important revision of these values has been published by Sackett et al. in 2022. A re-analysis of previous studies showed statistical biases in the form of overcorrecting for range restriction. After bias correction, the authors showed that while the relative order of the selection methods held up well, their absolute values of validity tended to be a bit lower than before – typically between 0.1 to 0.2 points lower on the validity coefficient.

Taking into consideration the revised estimates provided in Sackett et al. (2022), one of the most eminent predictors turns out to be a structured interview, going beyond cognitive ability tests. The latter showed a strong decline in capability for prediction in comparison with traditional standards. The revised estimates given in this particular study have a number of implications for those involved in practitioner roles. They state that firms should be realistic about a single selection procedure in regard to its prediction capability and focus more on a combination of prediction methods. They appear to be supporting systematic selection approaches but with somewhat lower prediction capability in comparison with traditional field theory.

In the current research, the results of the meta-analysis provide critical points of reference. They establish a degree of accuracy of prediction that may be meaningfully achieved in any selection model, including the ones that utilize ML techniques. They also validate the approach of multiple predictors over a solo predictor of a CV component. Finally, they emphasize the benefit of structured selection practices over unstructured ones – a change toward the very thing that the current research aims to make possible for the legal professionals' selection in the Ukrainian IT sector.

### 2.3 Machine Learning in Talent Acquisition

The application of ML to the process of employee selection is a natural outgrowth of the evidence-based practices that have been discussed earlier in this chapter. However, the employment of AI in human resource management has its own challenges in terms of the quality of data, transparency of the algorithm, and the process of change management in the organization (Tambe et al., 2019).

A framework for ML being effectively applicable to a certain task has been described by Brynjolfsson & Mitchell (2017). The factors that need to be taken into consideration for effective ML are: (1) the presence of a learned mapping between inputs and outputs; (2) sufficient example data available for training; (3) clearly determinable input variables; (4) mechanisms for correcting and improving the system; (5) stability in data with regard to time; and (6) sufficient data available for learning.

When these criteria are applied to the initial screening for CVs, there appears to be a high degree of applicability to ML. The inputs – the characteristics of the CVs – are already in a structured form or can be made to be; the output – the success or failure of a hire – is measurable; feedback can be gathered from follow-up information about performance; and the patterns of successful hiring should be relatively stable at least in a given context. The main challenge would appear to be the availability of thousands of legally hired examples against which patterns must be identified in most settings, necessitating approaches that can work with smaller samples.

The authors have a conception for "Suitability for Machine Learning" (SML) that serves as a measurable construct for making automation strategies. SML tasks have characteristics that make them amenable to algorithms: they entail repetition, involve well-structured information, and have the capability for verification of their output. Significantly in all of this is the point made by Brynjolfsson & Mitchell (2017) about the predominance of jobs being a combination of high SML tasks and low SML tasks. This suggests that the future of employment will be a process of human-machine collaboration.

This finding directly relates to the present research. The point of the research is not to substitute human judgment in the judicial selection of employees but to supplement the existing judgment. The initial screening of CVs involves a high-SML activity that can benefit from ML algorithms. Final interviews and selection involve low-SML activities that require the application of human judgment. The developed approach would serve as a triage system to allow the manager to spend their time on the candidates who have more to offer.

Also, it should be noted that the current research uses a hybrid approach that brings together traditional statistical modeling and analysis using LLMs. Statistical models aid in finding and validating particular predictors of candidate success, while LLMs examine unstructured narrative resumes. But this approach also capitalizes on the strong points of both methods: statistical models provide explanations for predictions made via explicit identification of underlying determinants for predictions, while LLMs provide better unstructured text pattern recognition. The hybrid screening tool validates statistical model predictions for candidate assessment and relies on LLMs for gaining insights from unstructured resumes.

### 2.4 Legal Sector Recruitment: Data-Driven Approaches

The report by JRG Partners (2024) adds to the existing literature a focus on candidate selection methods in the legal sphere. It concentrates on the rising importance of data-related candidate evaluation metrics including time to hire, quality of hire and retention measures. The use of data analytics in managing a talent pipeline is considered a means for better objectivity in candidate selection.

JRG Partners argues that a new paradigm for legal recruiting must be data-oriented and radically change the assessment of legal talent. Their report emphasizes that for such approach to be successful, time spent recruiting candidates and their subsequent performance should be carefully tracked. The argument put forward is that predictive analytics made possible through analyzing data generated based on previous recruitment decisions can help predict candidate success. This model also focuses on being more forward-thinking with regard to managing legal talent while utilizing external intelligence relating to compensation and candidate availability. By such techniques being put in place, legal recruiting would no longer be bound to credentials and contacts but would be capable of reaching a degree of objectivity similar to that found in other modern sectors.

The legal industry faces unique challenges in the application of data-intensive recruitment. Legal practices often cannot be measured effectively for the purposes of defining the boundaries of a 'successful' candidate. The inherently confidential nature of the law limits the pool of possible outcome data for the purposes of model training. Additionally, the smaller size of the average legal department may mean that many of organizations have insufficient historical candidate pool data for classic ML algorithms.

All such factors are compounded in in-house legal capacities in technology companies. While the efficiency of legal associates in a law firm can be measured fairly directly in billable hours, the same in case of in-house counsel is inherently harder to measure. This is because in-house counsel remediates situations that are inherently unpredictable, mitigates risks that may never come to fruition, and provides advice about which the full effect may only be apparent with the passage of time. Thus, a model for predictive hiring based on outcome definitions would face insurmountable difficulties because a successful in-house counsel would be hard to define until a number of years have passed.

Also, directly applicable to this research is the issue of transitioning from law firms to in-house lawyers. A great number of applicants for in-house legal positions in the technology industry come from a law firm background. Nonetheless, success in one environment does not equate to success in the other. The skill requirements vary drastically. Law firms require strong technical legal capability and a high level of billable hours, while in-house lawyers also need to be attuned to a business perspective with a focus on risk prevention and supplying actionable advice. Thus, a predictive model for recruiting would need to address such underlying distinctions instead of merely using law firm experiences as a simple indicator for success.

In this regard, a report by JRG Partners (2024) brings together scholarly research and needs within the industry to address legal recruitment using data. Specifically, their emphasis on overcoming bias while increasing efficiency resonates with the goal of this research effort – to design a screening instrument for Ukrainian IT legal managers. Finally, this report places legal hiring using data in a broader industrial setting that shows such a model to be theoretically valid but also indicative of current industrial trends regarding legal hiring.

### 2.5 Ukrainian IT Industry Context

The theories and empirical evidence introduced above are based on studies in developed Western economies. The application of those considerations in a Ukrainian environment requires a comprehension of a set of special characteristics for Ukraine's IT industry. In this context, a precious guide for a retrospective analysis of developments in talent management methods from traditional human resource management towards a more strategic workforce planning can be found in Cappelli & Keller (2014). Its importance in this capstone appears with regard to whether one should rely on intuition when selecting talents or instead follow systematic approaches based on facts. This remains even more critical for Ukraine because traditional methods of recruiting talents remain dominant here.

The pool of legal talent available to the IT industry in Ukraine corresponds to the country's educational system. Most of the lawyers have a degree in law obtained from a Ukrainian university but there is a growing trend of them having additional educational qualifications internationally. The proficiency in the English language varies significantly. There is a great need for a differentiation factor in terms of international engagement. The challenge caused by the lack of a general common law system calls for the lawyer to be trained separately to function well in a system of contracts/regulatory principles developed in common law countries.

The war that erupted in 2014 and escalated significantly in 2022 has also impacted the legal talent pool in numerous ways. Some of the lawyers have left the country, thereby contributing to a reduced talent pool. At the same time, a number of IT companies have shown strength and have continued to grow their headcount of legal talent. Even in the wake of remote work becoming the new norm, there may be challenges in the sense that new locations will have to be scouted for talent.

As already mentioned, the current position of Ukraine in the process of EU integration makes the landscape even more complex. Legal teams need to show their expertise in the GDPR framework, the consumer protection directives in the EU, and the rest of the regulatory frameworks in Europe. Such a context creates a need for lawyers who think internationally and continue to grow in knowledge – a skill set not reflected in traditional credentials but possibly identified in a detailed analysis of resumes.

Additionally, the competitive nature of the IT talent pool in Ukraine matters. Technology companies face not only domestic competition for qualified legal staff but also have to compete with international law firms and transnational corporations seeking to hire Ukrainian lawyers to staff their offices in Eastern Europe. Such competition will be relevant to both the number of qualified persons in the talent pool and the signaling aspect of prior work. Working for a distinguished company or law firm may be a signal of quality given the selective nature of the employer.

These contextual considerations imply that while general principles from more advanced markets may be informative, specific forecasts may require adjustment for the Ukrainian environment. Indicators of quality within Western labor markets may play out differently for Ukraine. The relative strength of various types of credentials, the validity and strength of various career paths, and the suitability factors for achievement within Ukrainian tech companies are empirical matters this research aims to investigate.

### 2.6 Literature Gap and Research Positioning

With this literature review I would like to highlight a substantial gap that is filled by this research. Despite considerable advancements from a theoretical point of view with respect to such topics as human capital, signaling, and person-environment fit, and more than several decades providing evidence for the validity of various selection methods and the increasingly important role of ML for recruitment processes, there is a significant lack of research on legal recruitment within the product-focused information and IT industry of Eastern Europe. To the best of my knowledge, there exist no previous studies focusing on the Ukrainian environment, within which local legal departments compete globally and under specific local conditions.

This research seeks to fill the gap with the first predictive model for screening legal positions within Ukrainian IT companies that is data-driven and specifically crafted for managerial application. It builds from proven bodies of research and tests the validity and application for a new set of circumstances. It uses proven methods and adjusts them to be able to adequately address the reduced sample size that comes with recruitment within specific professional categories.

The specific hypotheses that this research aims to test are the following:

(1) certain elements of the CV, namely education, experience, and language skills, can forecast recruitment success on the basis of human capital and signaling theories;

(2) indicators of fit informed by the narratives on the CV can offer predictive power on top of the other criteria available;

(3) an ML approach can offer similar validity to other approaches for recruitment; and

(4) a corresponding automation model to screen CVs can be created with the specific end-use considerations for legal managers in mind.

# CHAPTER 3. METHODOLOGY

## 3.1 Research Design: Mixed-Methods Approach

This research employs a mixed-methods design that aims to combine quantitative analysis of CVs criteria with a coding analysis of narrative content. Such approach is valid because legal recruitment success requires a number of quantifiable elements (e.g. years of prior work experience, tier of institution attended, language abilities) and qualitative elements (e.g. career goals, corporate culture fit) which numbers doesn't capture directly. This design is based on frameworks outlined in the previous chapter. The principles from Human Capital Theory apply to quantifiable features and Person-Environment Fit Theory – to qualitative analysis.

The quantitative aspect of this research entails the process of finding and extracting structural variables from CV and then analyzing the relationships between them and employment outcomes. Logistic regression analysis is being used as the primary method for identifying the predictive power of various features from CVs on employment success. As a complement to the method above, decision tree analysis is employed for detecting any possible non-linear associations and interaction effects that may be missed by linear analysis.

The qualitative part involves coding the narrative from CVs for extracting keywords relating to motivational and adaptational traits, career goal requirements, and elements concerning culture fit. Such coding procedure allows for structuring unorganized data elements in a manner amenable to use in a prediction model. This plays a critical role in assessing elements concerning a fit between individuals and the organization regarding career goals and values. These softer factors have substantial empirical support in being a determinant of retention and job satisfaction beyond those factors associated with credentials. They are central to understanding why some candidates succeed while others, with similar qualifications, do not.

This integration between the quantitative and qualitative aspects is carried out at the model development stage, wherein both variable types are incorporated within a customized screening tool based on GPT. This tool is essentially the end product derived from this research and is designated for utilization by legal managers for testing candidate CVs. The choice of the GPT platform was driven by its ability to process natural language as input and meet criteria for analysis involving statistics.

As for the research design, the study can be perceived as applied research with both explanatory and predictive objectives. In this case, the explanatory objective attempts to discover differences between successful and unsuccessful recruitment outcomes within the context of the IT industry and in-house legal function in Ukraine. On the other hand, the predictive objective aims at applying the research outcome to come up with a tool that would benefit future recruitment processes.

## 3.2 Data Collection

For this analysis, the data for this research was collected from two sources: a primary dataset containing legal professionals who were considered for hiring or hired by TechCo starting from 2017, and a secondary dataset from the LinkedIn group managed by me uniting Ukrainian in-house legal professionals and other specialists. In total, these datasets amount to 269 observations which can be analyzed for this research and represent legal professionals employed by or with interest in IT companies from Ukraine.

### *3.2.1 Primary Dataset: TechCo Candidates (n=178)*

The initial dataset is made up of 178 legal professionals in process of communicating with whom in the context of hiring to TechCo I have been personally involved since 2017. The significance of this dataset is that the information is holistic and spans the complete recruitment process from initial CV submissions to recruitment decisions and finally to their performances at TechCo for those who were hired. The longitudinal nature of this dataset, spanning more than seven years, allows for the evaluation of recruitment success formally at a stage after recruitment.

For each person within the primary dataset, the variables that can be assessed include:

(1) the original CV that was initially submitted as part of the application process;

(2) structured interview and bar raiser feedback notes;

(3) the outcome of the application process (whether an offer has been made);

and, for people who were hired:

(4) the amount of time they remained with the organization;

(5) performance feedback from periodic reviews; and

(6) current employment status (still with the organization, voluntary termination, or involuntary termination).

With this rich collection of data available for analysis, a number of indicators for "hiring success" can be created.

The first set consists of 178 participants with unique recruitment outcomes. From this number, there are 36 participants with successful recruitment outcomes; they were retained within the company for at least two years and achieved satisfactory or higher levels of performance. Another 12 participants were successfully recruited but either voluntarily left the company within two years or performed unsatisfactorily. The other 130 participants were not given a chance to be recruited and thus comprise a set for contrasting differences between successfully and unsuccessfully screened candidates.

The primary dataset has several advantages. Firstly, there is firsthand knowledge of all the cases being observed, allowing for coding of the variables with precision that might be otherwise subject to interpretations outside this context. Secondly, the observations were made within one context and so eliminate considerations that might be at play across multiple firms, such as differences within various companies. Thirdly, the outcomes achieved were well-verified and complete.

#### *3.2.2 Secondary Dataset: LinkedIn Professional Group (n=91)*

The data required for the secondary dataset, as stated earlier, was obtained from the professional networking group existing on LinkedIn for lawyers who have an interest in the field of law related to IT firms in Ukraine. As of the end of December 2025, this networking group had about 130 members, and from their profiles, which provide details not very different from resumes and can be used for analysis, I shortlisted 91 members.

The secondary dataset serves a number of purposes for the research. Firstly, it increases the sample size beyond what would be possible with the TechCo dataset only. Secondly, it adds a layer of variation across employers, allowing for a snapshot assessment as to whether any trends found within TechCo can be generalized to lawyers at other Ukrainian IT companies. Thirdly, this dataset represents the desired population for the screening tool, legal professionals actively engaged and seeking in-house IT roles, rather than a randomly selected sample from across all Ukrainian lawyers.

For each individual that is represented in the LinkedIn sample, the available data from a LinkedIn profile provides information that is similar to what is found in a CV: educational background, work experience, level of proficiency in a given language, licenses, and a biography. While it is not possible to directly equate a LinkedIn profile with a CV, it covers the same categories that the model is assessed upon. Another merit is evident with respect to the consistency of LinkedIn profiles.

A significant limitation of the secondary database is the lack of direct outcome measures, which include the success of hiring, performance, and ability to be retained. To overcome the limitation posed by the absence of direct measures, the study uses the proxy variable defined by the current employment situation and career advancement possibilities. For instance, people in senior legal positions in established information technology firms in Ukraine are considered to have background profiles that are “successful” in comparison to others in junior positions or others without jobs in the IT sector. Due to the application of this method, analysis is possible for the secondary dataset.

In cases where possible, additional outcome data for participants in the secondary sample was gathered through personal networks. For approximately 25 candidates at known career outcomes via industry networks, additional detailed success coding is done. These represent a validation sample for exploring whether findings from the primary sample hold for the broader population.

#### *3.2.3 Data Privacy and Ethical Considerations*

The use of personal information within the context of recruitment research raises important ethical and legal concerns, which for this research are incorporated through a complex of considerations and safeguards. The research process and design were created with close attention to current Ukrainian law and GDPR principles to take cognizance of the fact that this screening tool may be used by companies with EU operations as well as individuals within the EU (European Union, 2016).

In respect of the primary dataset, the necessary consent has been sought from current employees at TechCo whose data is included within the analysis. Where possible, consent has been sought from previous employees whose data is used within the analysis. In instances wherein such contact could not be made, the necessary anonymization has taken place prior to its inclusion within the dataset. It should be noted that such anonymization process removes any names and specific dates, though retains the specific variables for analysis. Additionally, all these datasets reside on a secured computer device accessible to me alone.

In the case of the secondary dataset, profile information was used for research purposes after being anonymized and aggregated. Also, the LinkedIn profile information used for the research was limited to information made public on the LinkedIn platform by members.

### 3.3 Feature Engineering: Extracting Predictive Variables from CVs

The process for feature engineering herein is informed by the overall framework from the literature review and is guided by Human Capital Theory for credential-based variable creation and Person-Environment Fit Theory for environment/fit variable creation.

#### *3.3.1 Quantitative Variables*

The quantitative variables were gathered from CVs through coding with respect to objective qualifications and indicators of experience. The quantitative variables take into consideration the Human Capital Theory that points out educational investment and development, as well as the Signal Theory concentrating on visible qualifications which signal unobservable ability. Table 1 summarizes the complete list of quantitative variables gathered from the CVs.

**Table 1**. *Quantitative Variables Extracted from CVs*

| Variable Category | Variable Name | Operationalization |
|---|---|---|
| Education | University Tier | 1=Top 5 UA law schools, 2=Other UA, 3=International |
| | Degree Level | 1=Bachelor, 2=Specialist/Master, 3=LLM/PhD |
| | International Education | Binary: Any degree/certificate from non-UA institution |
| Experience | Total Years Legal | Continuous: Years since first legal position |
| | IT Company Experience | Binary: Any prior in-house IT company role |
| | Law Firm Experience | Binary: Any prior law firm role |
| | Big 4 / International Firm | Binary: Experience at Big 4 or international firm |
| | Number of Employers | Count: Total distinct employers listed |
| | Average Tenure | Continuous: Mean months per employer |
| Language | English Level | 1=Basic, 2=Intermediate, 3=Upper-Intermediate, 4=Advanced, 5=Native |
| | English Certification | Binary: IELTS/TOEFL/Cambridge certificate listed |
| | Additional Languages | Count: Languages beyond Ukrainian/English |
| Certifications | Bar Admission | Binary: Ukrainian advocate license |
| | International Qualification | Binary: Non-UA bar/solicitor qualification |
| | Specialized Certifications | Count: CIPP, CIPM or similar credentials |
| Practice Areas | IP/IT Experience | Binary: IP, IT or data protection listed |

| Variable Category | Variable Name | Operationalization |
|---|---|---|
| | Corporate/M&A Experience | Binary: Corporate, M&A or venture listed |
| | International Transactions | Binary: Cross-border work explicitly mentioned |

*Note. Variables are coded from CV and LinkedIn profile data. Tier 1 Ukrainian universities include Taras Shevchenko National University of Kyiv, Ivan Franko National University of Lviv, Yaroslav Mudryi National Law University, and comparable institutions.*

The coding system for quantitative variables strictly follows specific criteria to promote consistency across the CVs. In the case of variables involving evaluative judgment (such as English proficiency when there is no actual scoring system available), coding utilizes contextual cues: candidates with evidence of negotiating international contracts or having studied overseas would be assigned higher codes than those without such cues. For situations involving ambiguity and missing data from the CVs, a conservative coding strategy is utilized (for example, coding English proficiency at a lower level when there is ambiguity).

### *3.3.2 Qualitative Variables*

Thematic coding of unstructured parts of a CV, particular to job descriptions, career statements, and available cover letters (when they are integrated as part of a CV), is used to produce variables for qualitative indicators such as values fit, career/personal motivation, and workplace interests. This serves a critical function in Person-Environment Theory because assessing such factors determines whether a candidate would be a good fit for a workplace with a particular culture. Literature has shown that interpersonal and contextual skills appraisal (whether direct or indirect) creates incremental validity in predictions of academic achievement in addition to job performance when using cognitive and credential-based predictors (Lievens & Sackett, 2012). By using this coding in variables for the decision-making process of the GPT screening tool, a candidate's hard and soft qualifications can be considered.

The coding process was carried out based on the iterative analysis of the TechCo dataset. Initial categories were derived from open coding of the narratives from CVs and were then consolidated within a structural codebook. Table 2 summarizes the qualitative variables and their operationalizations.

**Table 2**. *Qualitative Variables Extracted from CV Narratives*

| Theme Category | Variable Name | Coding Indicators |
|---|---|---|
| Career Orientation | Tech Industry Interest | Descriptions involving technology and innovation; involvement with a personal project/hobby relating to technology |
| | Business Orientation | Business-oriented language with a focus on results, commercial awareness and analysis |
| | Growth Mindset | Matters of education, professional development, and engagement with new tasks, with a focus on improvement |
| Work Style | Adaptability | Examples of role transformation, changes within a particular industry, and ability to handle ambiguity with adaptability |
| | Proactivity | Language that conveys initiative, self-direction, and a predisposition towards creation as opposed to followership |
| | Team Collaboration | Mentioning cross-functional teamwork and collaboration, with a focus on supporting others |
| Achievement Focus | Quantified Results | Including specific quantitative indicators such as numbers, percentages, and other measurable results in describing professional activity |
| | Leadership Evidence | Background in managing staff, running projects and being responsible for their results |

| Theme Category | Variable Name | Coding Indicators |
|---|---|---|
| | Problem-Solving Examples | Identification of specific problems solved, solutions developed, and hurdles overcome |
| Cultural Indicators | International Exposure | Practical experiences with travel, foreign assignments, multicultural environments and a global perspective |
| | Entrepreneurial Signals | Engagement in side projects, entrepreneurship, startup involvement, and risk tolerance assessment |
| | Values Alignment | Evidence of alignment with IT/start-up culture values such as swift execution, innovation, and measurable results |

*Note. The variables are measured on a scale of three (0, 2), where 0 signifies the lack of it, 1 signifies the limited presence, and 2 signifies the presence of it. This scale is developed using the presence of the prominent indicators in the resume.*

For each qualitative variable, coding was done on a three-point scale: 0 (absent), 1 (weakly present), and 2 (strongly present). The design aims at finding a middle ground between the necessity for more refined classification and the limitations that come with subjective evaluation. Inter-rater reliability was tested for all the variables on a sample of 20 randomly selected CVs with Cohen's kappa above 0.7 (Cohen, 1960).

These qualitative variables allow for the identification of factors predicted to be correlated with person-organization fit from prior literature. Tech Industry Interest and Business Orientation factors represent alignment for the tech industry context; while Adaptability and Proactivity represent alignment for the dynamic and uncertain environment found within growth-stage companies. International Exposure and Entrepreneurial Signals refer to alignment with the organizational context at TechCo, including the start-up environment and the international business. These factors allow for further alignment with candidates' qualifications and representation outside of what was quantified.

### 3.4 Model Development Approach

#### *3.4.1 Platform Selection: Custom GPT vs. Alternatives*

In the research, several platform alternatives were utilized for the application of the screening model. These platforms include traditional machine learning approaches implemented using a customized application, a fine-tuned LLM approach, and a customized GPT model with proper instructions. Every platform alternative has its strengths and limitations.

A customized GPT was found to be the best suitable platform for this research. The GPT platform developed by OpenAI is utilized for creating specific models that can harness the natural language processing power of GPT-5.2 while incorporating domain-specific instructions within them. This approach brings several benefits with it: it can process unstructured CV text, and there is minimal necessity for any infrastructural setup other than a subscription to ChatGPT without depending on any support team for any data-related services. Additionally, this approach can generate natural language-based explanations for any evaluation made.

The GPT custom model is consistent with the framework for proper application of ML as outlined by Brynjolfsson & Mitchell (2017). In particular, the screening process is classified as a high-SML task and is well-suited for algorithmic support while the GPT system retains human intervention. Additionally, the system provides feedback and explanations meant to support, and not replace, human decision-making. This is significant with regard to bias within the recruitment process carried out by algorithms.

#### *3.4.2 Training Data Structuring*

In a custom GPT model, there is no training process as is the case with traditional ML. The detailed instructions cover the patterns derived from the analysis process. The concept of training data represents the knowledge base that is used to inform the GPT on its evaluation criteria. This is collected from the primary and secondary datasets based on the process explained below.

To begin with, the process uses logistic regression analysis to identify the quantitative and qualitative variables with significant predictive ability for hiring success within the TechCo sample. A p-value $<0.10$ is taken as the criterion for retaining these significant variables for inclusion in the screening model for hiring success, balancing sensitivity with avoiding over-inclusion of marginal predictors. The strength and direction of the coefficients correspond to the influence and effect on the screening process.

Second, decision-tree analysis identifies threshold and interaction effects that may be hidden from the logistic model. For example, if IT Company Experience is significant only for candidates with high English proficiency, then this interaction is incorporated within the GPT instructions as a separate decision rule. Decision-tree analysis is useful for identifying the criteria employed by the GPT system for filtering candidates based on their CVs.

Thirdly, profiling for successful recruits provides prototype descriptions that function as benchmarks for the GPT. The descriptions identify typical traits possessed by lawyers who have been successful at TechCo and thus allow the GPT to determine the similarity between potential candidates and previous successful recruits. The profiles created here are anonymized and aggregation-based and therefore do not identify specific individuals.

This structured knowledge base is then synthesized as a detailed instructional document that is incorporated within the system prompt for the GPT. This document spells out: (1) the criteria that needs to be assessed for various CVs, (2) the considerations for scoring each criteria parameter, (3) the priorities given to various considerations for the overall assessment task, (4) the threshold conditions for receiving specific recommendations, and (5) sample assessments for proper application of considerations. The instructional document is iteratively refined through testing with sample CVs held out for this purpose.

#### *3.4.3 Prompt Engineering and Instruction Design*

Prompt engineering can be described as the design and construction of instructional material for governing GPT responses and is a core methodology for application that determines the efficiency with which the tool applies research findings. Design principles for prompt engineering essentially include best practices for developing personal GPT solutions while incorporating necessary domain knowledge for legal recruitment evaluation.

The GPT system prompt is divided into several parts. The first part defines the role and task for the assistant: screening legal candidates for IT companies, seeking candidates for in-house recruitment within the Ukrainian IT industry. The second part defines the list of quantitative indicators that should be extracted from the CVs sent by the candidates. The third part defines the indicators to be examined and gives definitions for each one with examples for strong and weak indicators. The fourth part discusses the evaluation algorithm and explains how the scoring for each variable is calculated to come to a cumulative suitability score. This comes directly from the regression coefficients and decision tree produced from the analysis. The fifth part discusses guidelines for model output and requires the GPT to produce both a suitability score between 0-100 and a text analysis explaining a candidate's positives and negatives. The sixth part defines ethical parameters and prevents suitability analysis from being carried out on protected categories.

The prompt requires the use of few-shot examples that come with sample CVs and model judgments on the evaluation criteria. These examples can help calibrate the GPT model's evaluation judgment on the criteria being used. They are drawn from the primary dataset and represent a spectrum of outcomes ranging from unquestionable positives and negatives to borderline positives.

The output format is designed for managerial usability. The GPT creates a formal evaluation consisting of: (1) overall suitability score and explanation (for instance: "78/100 – Strong candidate, proceed with interview"); (2) key evaluation scores (for qualifications, alignment with experience, cultural fit); (3) description of personal strengths found within the resume; (4) analysis of potential shortcomings; and (5) finally, a recommendation (to proceed with the interview, additional screening is necessary, and so forth). It is beneficial for the managers recruiting candidates, providing them with crucial details while simultaneously being transparent about their evaluation criteria.

The complete GPT system prompt is provided in Appendix B. This prompt translates the statistical model into executable instructions for the LLM-based screening tool.

### 3.5 Validation Strategy

The need to establish criterion-related validity requires robust methods and strategies in the process of validation to avoid the challenges and pitfalls associated with the methodology, such as restriction in range, contamination of the criterion, and issues associated with fit (Van Iddekinge & Ployhart, 2008). Various methods in the current study are implemented to validate the screening model and ensure true prediction and no sample-based artifact in prediction validity.

Validation of the screening model is directed at two inquiries: first, whether the model accurately reproduces the hiring outcomes observed in the sample on which it was developed, and second, whether the model can be generalized to new candidates on whom the model was never trained. Different methods can be used for each.

Internal validity is assessed on the principal TechCo data set by means of cross-validation. Because the sample is small, leave-one-out cross-validation (LOOCV) is employed: the model is trained on all instances except one held out example and predicts on the held-out example, repeating for each instance in the data set. This ensures maximum use of the training data with each repetition and produces unbiased estimates for evaluation.

Model evaluation criteria include classification accuracy (proportion of correct predictions), sensitivity (rate at which successful candidates were predicted with actual success), specificity (rate at which unsuccessful candidates were predicted with actual failure), and the AUC for the receiver operating characteristic (Hanley & McNeil, 1982). The AUC is highly informative about the model's ability to distinguish between successful and unsuccessful candidates for all possible threshold values. Based on the criteria suggested by Schmidt & Hunter (1998) and later supported by criteria suggested by Sackett et al. (2022), AUC between 0.65 and 0.75 represent values for moderate predictive validity similar to current best methods of selection.

External validity was assessed on a secondary dataset. The TechCo model was tested on the set of LinkedIn group members for whom outcome information was gathered. Agreement on internal and external validity tests would confirm model generalizability on other Ukrainian IT companies; otherwise, company-specific adaptation would be required.

The GPT then undergoes additional validation checks to ensure proper alignment with the statistical model. A sample of CVs is tested for alignment between the statistical model (manual variable assignment and then the regression equation) and the GPT. If there is agreement between the two approaches, this ensures the prompt design has properly captured the analysis results. A lack of agreement triggers adjustments to the prompt design until satisfactory agreement is achieved at a rate above a correlation of 0.85.

Future validation would be pursued through deployment validation. As the tool is used for screening candidates at TechCo, the predicted outcomes can be compared with the actual final employment decision taken. Future validation would be done with periodic updates to the model based on accumulated prospective validation data. The research design would allow for periodic updates to the model.

### 3.6 Limitations of Methodology

The methodology described within this chapter suffers from several drawbacks which should be taken into consideration and impact the discussion surrounding the findings.

The biggest constraint is represented by the sample size. With a sample size of 178 observations from the prime dataset, the number of predictor variables that can be estimated is limited. Traditional best practices state that for the application of logistic regression analysis of at least 10-15 observations should be available for each predictor variable. With around 20 predictor candidates available, this analysis needs to identify the most relevant predictor on the basis of qualitative knowledge and may lack sufficient power for detecting small effects.

The fact that the focus is on one particular organization leads to the issue of generalizability being raised. This is mainly because predictive factors for being successful at TechCo may not be applicable to other IT enterprises in Ukraine with cultures and sizes among other differing facets. Despite the advantage presented by the secondary dataset being available for generalizability analysis, the fact that there is unverified outcome information for the bulk of the members within the secondary dataset is a limiting factor towards external validation.

A bias may arise from the involvement of the researcher in recruitment and model building. In evaluating the candidates' success rates, my personal perceptions about the qualities for a good lawyer may influence the process and be incorporated at both stages. However, by using objective criteria (such as tenure and actual performance ratings) for determining the outcome definitions, this problem can be reduced. Also, replication by unconnected researcher would increase the validity of the study.

Outcome definition also raises several difficulties. In the current research endeavor, the concept "hiring success" is adequately captured through a combination of tenures, performance measures, and employment status. Though, these may not be the best means for measuring the value contributed by a lawyer. In this context, it should be clarified that a lawyer having worked for 18 months and then resigning for a superior opportunity would be regarded as less successful than another lawyer having worked for three years with merely satisfactory performance.

Temporal factors might play a role with respect to validity. The sample includes recruitment data spanning from 2017 to 2025. This time range encompasses a dramatic shift in the Ukrainian information and legal sectors. When predictive trends were informative and helpful for the year 2018, they may no longer be relevant by the year 2025. The year 2022 and the resulting full-scale war add complexity to this analysis environment, wherein dramatic changes have taken place within candidate pools and companies.

Additionally, the GPT system introduces other sources of uncertainty. While the custom GPT provides benefits and is useful, with regard to transparency and structure as a statistical model and intended assessment logic, it can still be categorized as not fully transparent. The validation process involving comparison between the GPT response and the prediction made by the statistics model ensures consistency; however, full consistency may not be achieved. Additionally, the GPT system keeps on changing with updates from OpenAI to their platform.

Despite these difficulties, this approach represents a rigorous method for developing a useful tool under real-world conditions. The mixed methods design, guideline-driven variable inclusion, variety of validation methods utilized, acknowledgment and proper caveats for conditions. All of them add to overall confidence within this screening tool while adjusting for the strength of claims being made.

# CHAPTER 4. DEVELOPMENT OF THE PREDICTIVE MODEL

## 4.1 Variable Identification and Operationalization

The chapter on methodology found a total of 18 quantitative and 12 qualitative variables from the CVs. However, not all of the identified variables possess predictive power on the measure of recruitment success within the TechCo environment. This chapter discusses the application of the conversion process for identifying the predictive variables for inclusion within the screening system.

Logistic regression analysis was undertaken on the primary TechCo dataset (n = 178), with the dependent variable being the outcome measure for "Hiring Success" (as described subchapter 4.2). In the initial model including all candidate predictor variables, there was evidence of both multi-collinearity and overfitting. A stepwise approach was employed and produced a simplified model with satisfactory model fitting qualities (Nagelkerke $R^2 = 0.47$, Hosmer-Lemeshow $p = 0.38$; Hosmer & Lemeshow, 2000) and maintained all predictor terms with p-values $< .10$. Table 3 shows the multivariable regression model.

**Table 3**. *Logistic Regression Results: Predictors of Hiring Success*

| Variable | B | SE | Odds Ratio | p-value |
|---|---|---|---|---|
| English Level | 0.89 | 0.31 | 2.44 | .004** |
| IT Company Experience | 1.42 | 0.48 | 4.14 | .003** |
| Average Tenure (months) | 0.03 | 0.01 | 1.03 | .012* |
| Tech Industry Interest | 0.76 | 0.29 | 2.14 | .009** |
| Business Orientation | 0.68 | 0.27 | 1.97 | .012* |
| Adaptability | 0.61 | 0.25 | 1.84 | .015* |
| International Transactions | 0.94 | 0.39 | 2.56 | .016* |
| University Tier | 0.52 | 0.28 | 1.68 | .063† |
| Growth Mindset | 0.49 | 0.26 | 1.63 | .059† |
| Constant | -4.82 | 1.14 | 0.01 | <.001*** |

*Note. N = 178. Nagelkerke $R^2 = .47$. †$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$*

On this basis, the final screening model included nine predictor variables with differential weights. The operationalization framework transformed the regression coefficients into a scoring system with a range of 100 points. The system used the odds ratio for each variable as the weight. Table 4 highlights the scoring system used within the GPT decision logic.

**Table 4**. *Scoring Weights for Predictive Variables in GPT Assessment*

| Variable | Maximum Points | Scoring Rule |
|---|---|---|
| English Level | 20 | 4 pts per level (1-5 scale) |
| IT Company Experience | 18 | 18 if present, 0 if absent |
| International Transactions | 12 | 12 if present, 0 if absent |
| Tech Industry Interest | 12 | 6 pts per level (0-2 scale) |
| Business Orientation | 10 | 5 pts per level (0-2 scale) |
| Adaptability | 10 | 5 pts per level (0-2 scale) |
| Average Tenure | 8 | 2 pts per 12 months avg, max 8 |
| University Tier | 6 | 6=Tier1/Int'l, 3=Tier2, 0=Other |

| Variable | Maximum Points | Scoring Rule |
|---|---|---|
| Growth Mindset | 4 | 2 pts per level (0-2 scale) |
| **Total Maximum** | **100** | |

*Note. Weights are proportional to odds ratios from logistic regression, scaled to sum to 100.*

The complete data collection instruments, including coding forms and decision rules for each variable, are documented in Appendix A.

**4.2 Definition of Success Criteria: Criteria for a "Good Hire"**

The validity of any predictive tool for recruitment is dependent on the definition given to "success." A predictive tool may be highly optimized for one definition of "success" and suboptimal for other definitions. This section outlines the criteria for "success" used to validate and develop the screening model and discusses the difficulties and complexities associated with measuring "success" within the legal profession.

The first criterion for measuring success within this research is a compound one that includes three aspects: approval for employment, retention rate, and overall performance appraisal outcomes. A candidate is regarded as "successful" if they meet all the following criteria: (1) they were at least granted a position offer (for candidates at this stage); (2) they were retained at the company for at least 24 months or are currently employed; and (3) they were rated "meets expectations" or higher on their periodic performance appraisals.

The retention threshold at 24 months was arrived at by examining the recruitment statistics for TechCo and finding that the majority of the attrition happens within the first two years for the company. Employees who survive this length of time demonstrate enough loyalty and integration to be considered a success from a retention perspective. It is further optimized for pragmatic purposes, and any shorter time would be overly optimistic about the retention levels and would drastically reduce the number of employees qualified for retention analysis.

The critical evaluation at TechCo is done on a scale of five: (1) Does Not Meet Expectations, (2) Partially Meets Expectations, (3) Meets Expectations, (4) Exceeds Expectations, and (5) Significantly Exceeds Expectations. In regard to a classification of succeeding, a score of 3 and above is taken as a criterion for succeeding. A more relaxed criterion was used here as opposed to a stricter one such as requiring "Exceeds Expectations" (4+) for classification as succeeding. The basic intention for screening is to eliminate the unsuccessful and not merely to identify the highly exceptional ones.

For the candidates that were not given offers, categorizing them required a different strategy. These candidates cannot be said to be "unsuccessful" since they were never given the chance to perform at TechCo. However, this process is a form of evaluation for potential for success. For candidates that made it to the final round for interviews and were not given offers, they were categorized as "unsuccessful” candidates for the model evaluation while recognizing that this is the joint decision of the recruitment committee and not a failure in actual performance.

The definition of success was tested for validity with a sensitivity analysis. Alternative definitions were tested: (1) retention-only (any candidate still employed regardless of success); (2) performance-only (any candidate with a rating of 4 or higher regardless of time); and (3) strict composite (candidates with at least 24 months tenure plus a rating of 4 or higher). The overall composite definition showed the best compromise between discrimination and reliability with AUC =0.74 (vs. AUC =0.68 for retention-only and AUC =0.71 for performance-only criteria). The strict criteria showed higher discrimination with AUC =0.78 but with much reduced numbers for the success class (only 26 candidates remained).

When these criteria are then applied to the initial dataset (n = 178), the classification can be achieved as follows. 48 candidates can be identified as being successful (27%), including 36 candidates selected on the criteria of retention and performance and 12 candidates selected on the basis that they were given offers and refused (classified as being successful on having achieved the offer stage). One hundred thirty candidates can be identified as being unsuccessful (73%), including nine candidates selected on the basis they were hired but either remained for a short time and/or underperformed and candidates qualified for having passed various stages prior to being denied offers at later stages (a total of 121 candidates). A baseline rate for classification exists at 27% and therefore represents a class-imbalance issue that is corrected for and handled appropriately with necessary advanced statistical analysis.

### 4.3 Dataset Preparation and Anonymization

Before any model building was carried out, the raw data underwent some preparation tasks with the aim of ensuring quality and completeness as well as anonymizing the information. This subchapter highlights the various tasks undertaken to ensure that the created dataset is ready for analysis.

Data cleaning incorporated several aspects that were noticeable within the raw CV dataset. Missing values were a problem for around 8% of observations, particularly within the categories requiring inference from limited information (such as English Level for candidates for whom no proficiency claim was made). Missing values were dealt with by adopting conservative methods for imputation. Categorical datasets were assigned the lowest relevant category and continuous datasets were replaced with the sample median for the particular variable. Sensitivity analysis showed that the findings were immune to other ways of treating this issue.

Both variable transformation and multiple predictor variables were utilized. Experience was transformed by taking the natural logarithm to normalize the positively skewed variable. Average tenure was winsorized at the $95^{th}$ percentile to remove the effect of outliers, namely candidates with very long service with one company. Qualitative research was standardized on a scale from 0 to 2 as explained under methodology.

To ensure the privacy of individuals and to allow for analysis, anonymization framework was used. The protocol included three stages. Stage 1 suppressed direct identifying information: names were replaced with sequential codes (C001, C002, and so on); email addresses, telephone numbers, and street addresses were removed; profile images were not included in the analysis. Stage 2 transformed quasi-identifiers: names of the institutions were replaced with tier codes; names of the employers were replaced with categorical codes (for example, "Big 4," "Ukrainian IT," "International Law Firm"); specific types were transformed to duration features (year and month). Stage 3 suppressed attribute combinations: observations with distinctive combinations of features were selected as potential for re-identification were discovered; for said observations with distinctive features, the least important attribute was further generalized.

Validation of anonymization was carried out through re-identification attempts. A third-party reviewer with prior knowledge of the Ukrainian legal market was used for attempts at re-identifying people from the anonymized dataset. The reviewer was able to successfully identify 3 out of 178 (1.7%) subjects with relative ease; all were highly public figures with career backgrounds well documented. For this set, additional generalization was carried out to make re-identification impossible. On all subjects, k-anonymity with k = 5 was achieved for the anonymized dataset. This meant that all subjects share their set of quasi-identifiers with at least four other subjects.

The anonymized dataset was organized under two formats for two purposes. The analysis dataset retained all the transformed variables within a spreadsheet format amenable to analysis with statistical software. The GPT training dataset was reconstructed for a combination of anonymized case profiles with their respective success designations for integration with the GPT system prompt as a set of examples for the model to learn from. Table 5 below shows a sample anonymized case profile for training the GPT model.

**Table 5**. *Example Anonymized Case Profile for GPT Training*

| Attribute | Value |
|---|---|
| Case ID | C047 |
| Education | Tier 1 Ukrainian university, Master's degree |
| Experience Summary | 6 years total; 3 years at Ukrainian IT company; 2 years at international law firm; 1 year at local firm |
| English Level | Advanced (Level 4) |
| Practice Areas | Corporate, M&A, cross-border transactions |
| Narrative Indicators | Strong tech interest (references startup ecosystem); business-oriented language; evidence of adaptability (multiple role changes) |

| Attribute | Value |
|---|---|
| Average Tenure | 24 months |
| Outcome | SUCCESSFUL (3+ years tenure, Exceeds Expectations) |

*Note. Profile shown is a composite example, not an actual individual from the dataset.*

### 4.4 Building the GPT-Based Screening Tool

The construction and use of the GPT screening system accomplishes the application of the statistics and case profiles to a system for screening candidate CVs. This subchapter discusses system design aspects and generation parameters for deployment within the screening system.

#### *4.4.1 System Architecture*

The screening tool was designed as a customized GPT developed on the basis of the OpenAI GPT-5.2 platform (OpenAI, 2025) using the ChatGPT Plus interface. The structure for the customized GPT involves three aspects: the system prompt that holds evaluation instructions, the knowledge repository supporting reference profiles, and the conversation interface that is utilized for submitting CVs and generating evaluation feedback.

The system prompt is the primary controlling mechanism and houses the logic for evaluation as inferred from the analysis and is composed of about 3,500 tokens. The system prompt is designed to offer all the necessary information within the constraints of the context window and lasts throughout all interactions with the customized version of GPT. This ensures that the criteria for evaluation apply regardless of the number of CVs being assessed.

The knowledge base consists of anonymized case profiles with a total of 15 instances for capturing the range of observed outcomes from the training cases: specifically for this research, there were instances classified as strictly successful cases with 5 instances, strictly unsuccessful cases with 5 instances, and borderline cases with the other 5 instances. The instances were used as a form of a few-shot learning example to train the judgment for the GPT. These instances utilize the system prompt and do not take the form of uploaded files. The configurations for each instance are determined in Table 5.

The interface supports natural language interactions between the user and the tool. A user can submit a resume through text paste, PDF file upload, as well as providing a link connecting to LinkedIn profiles (if publicly available). The GPT model identifies relevant information from any given resume structure and provides a resume analysis based on the logic applied to the model. The interface is designed to allow for further questions that would give insight to the candidate on specific aspects assessed and for additional information on a candidate.

In terms of privacy protection features, these were incorporated into the design structure. The design for the custom GPT is such that the conversation history is not retained between discrete sessions. This ensures that candidate information is not accumulated. The system prompt is specifically designed such that the GPT does not reference any information from previous candidates while screening the current ones.

#### *4.4.2 Instruction Set Design*

The instruction set is the heart of the system prompt and defines the process whereby the GPT evaluates CVs. The instruction set was created through a process of refinement from a straightforward conversion from the statistical model and then tested on known examples. The finished instruction set comprises six sections, each with a particular function to perform within the evaluation process.

Section 1: Role Definition defines the role of the GPT and its purpose. The task description fixes the role of the assistant as a screening device for legal candidates hired by globally focused Ukrainian IT companies. The definition of role specifically mentions that the device provides suggestions for human decision-making and does not make the final decision on recruitment.

Section 2: Information Extraction defines the information to be gleaned from each CV. The guidelines offer definitions for each one of the nine predictive factors and methods for handling uncertain cases. For example, English Level is assessed on the basis of clear indicators (for example, the ‘Advanced English’ wording), contextual information (such as overseas educational background and experience with overseas clients), and the quality of English expressed in the CV if written in English. This information-gathering process is intended to facilitate reliable coding with respect to a broad variety of CV forms.

Section 3: Scoring Algorithm is utilized for quantifying the extracted information. The guidelines identify the number of points allotted for every variable as listed under Table 4. The algorithm uses threshold checks that were created from decision tree analysis: candidates with a total score below 40 automatically receive "Not Recommended" regardless of other criteria, while candidates with English Level lower than 3 are flagged for this requirement.

Section 4: Qualitative Assessment provides criteria for qualitative assessment on aspects of the story that cannot be accurately assessed by numerical ratings. The criteria inform the GPT on aspects to evaluate for overall story cohesiveness and content that lacks alignment between goals and experiences and red flags such as gaps and negative commentary about previous employers. Findings from qualitative analysis can be incorporated within the story evaluation and cannot be incorporated within the numerical evaluation.

Section 5: Reference Profiles defines the list of the 15 anonymized case descriptions and their respective evaluation. Every reference profile comprises a brief description of the candidate's background information, the number of points allotted for each variable tested, the overall score achieved, whether the application is declared successful/unsuccessful, and a reason explained for this specific evaluation. The GPT is programmed to apply the reference profiles for validation: when testing a fresh CV, similar reference profiles must be found for consistency with previous assessments.

Section 6: Ethical Constraints outlines impermissible behaviors and protected features. The guidelines strictly prohibit the use of gender, age, marital status, disability, ethnic origins, religion, and political convictions as factors for producing any kind of assessment. In this regard, the system is required to flag and disregard any such details found within the uploaded CV instead of allowing these factors to be included within the assessment results. In addition to this consideration, the constraints strictly prohibit the production of recommendations on the appearance of any candidate when personal photographs are included and must state limitations when there is inadequate information for a reliable assessment.

#### *4.4.3 Output Configuration (Percentage Score / Binary Recommendation)*

The output structure determines the way and form through which evaluation outcomes are communicated to users. The structure strikes a balance between being comprehensive and providing sufficient information for proper decision-making while being usable and preventing information overload. The structure consists of three components: summary score, detailed report, and narrative assessment.

The summary score provides a summary suitability score on a scale from 0-100 with a corresponding categorical recommendation. The score range is mapped to the categories as follows: between 70 and 100 inclusive: "Strongly Recommend Interview"; between 55 and 69 inclusive: "Recommend Interview"; between 40 and 54 inclusive: "Consider with Caution"; below 40: "Do Not Recommend". These particular threshold settings maximize both sensitivity and specificity for predicting whether a candidate succeeds on the basis of "Recommend" at the 55-point threshold: 82% would pass successfully and rule out 71% of failures.

A more detailed explanation lists out the scores for each of the nine predictive factors to facilitate an appreciation for the considerations included within the overall score. In doing so, several objectives can be realized: the ability for the user to determine whether they can associate the overall score with their perception from the resume; pointing out areas of strength and weakness for analysis during later interviews; and greater trust within the model by no longer treating the decision as a black box for the recommendation made.

The narrative assessment provides a brief text analysis with approximately 200-300 words that summarizes important observations, identifies key strengths and possible areas of concern, and explains the reasoning for the general recommendation made. A general structure for the written analysis will be to: offer a one-sentence summary response; list two to three bullet points for important strengths; list one to two bullet points for potential areas of concern; and finally close with a summary statement and suggested content for potential interview emphasis should the process continue. A sample result for a fictional candidate can be seen within Figure 1 below.

**Figure 1**. *Example GPT Screening Tool Output*

**CANDIDATE ASSESSMENT SUMMARY**

**Overall Score: 72/100 – STRONGLY RECOMMEND INTERVIEW**

**Component Scores:**

• English Level: 16/20 (Advanced)

• IT Company Experience: 18/18 (Present)

• International Transactions: 12/12 (Present)

• Tech Industry Interest: 8/12 (Moderate)

• Business Orientation: 8/10 (Strong)

• Adaptability: 5/10 (Moderate)

• Average Tenure: 4/8 (18 months average)

• University Tier: 3/6 (Tier 2)

• Growth Mindset: 2/4 (Some evidence)

**Narrative Assessment:**

This candidate presents a strong profile for an in-house legal role at a Ukrainian IT company.

**Key Strengths:**

• Prior IT company experience provides directly relevant context

• Cross-border transaction experience aligns with international operations

• Business-oriented language suggests commercial awareness

**Areas for Interview Focus:**

• Relatively short average tenure (18 months) – explore reasons for transitions

• Limited explicit tech enthusiasm in narrative – assess genuine interest

*Note. Example output is illustrative and based on a hypothetical candidate profile.*

The tool produces structured assessments with numeric scores and narrative explanations. Representative examples across different score ranges are provided in Appendix C, illustrating how the tool evaluates candidates with varying qualifications and experience profiles.

The output also comes with a binary recommendation system that is meant for users who want simple decision-making. The system can be triggered through a command such as "Give me the short version," and the GPT responds with just the recommendation category (Recommend / Do Not Recommend) and a one-sentence explanation. This is for screening purposes for a high volume of candidates for which detailed analysis may be unfeasible.

### 4.5 Pilot Testing and Iterative Refinement

Before actual deployment, substantial pilot testing was carried out on the screening system with respect to evaluation for potential deficiencies and improvements within the realms of evaluation accuracy and system output. The pilot testing process was carried out with three phases: validation on the training dataset, validation on the holdout sample, and user acceptance testing with legal managers.

The purpose of phase one internal validation was to determine whether the GPT system was able to replicate the statistical model. A sample of thirty CVs was selected from the training sample and was assessed under both approaches: the statistical model (manual analysis calculated with the regression equation), and the GPT system. The initial findings suggested that there was a considerable disparity between the two approaches, with the GPT system generating higher scores than the statistical model for twenty-three out of thirty candidates, with a mean disparity of 8.3. A closer review suggested that this was due to the word choice used within the evaluation criteria that allowed the GPT system to assign more relaxed interpretations than originally planned due to 'equivocal' wording.

The next stage was the refinement process after Phase 1. It included the upgrading of the instruction set to provide more specific advice on scoring. The revised instructions included specific illustrations of evidence that would meet criteria for a given scoring level, specific advice on being conservative in coding for instances when evidence was ambivalent, along with numerical anchoring that suggested the median score on the reference population was at 45 points. Further re-testing showed a mean scoring discrepancy of 2.1 with a correlation of 0.91 between GPT and the statistical model-based scores.

The Phase 2 external validation process tested the ability of the model to perform well on observations that were not used to train the model. The hold out sample included 25 observations selected from the second dataset for whom outcome data was available through professional networks. The GPT was then blindly tested on the CVs for predicting success outcomes to be compared with actual success designations. It achieved a sensitivity and specificity rate of 0.69 and 0.75 respectively with actual predictive values amounting to 18 out of a total 25 (72%), and the AUC was found to be moderate at 0.73. It identified correctly 9 of 13 successful cases and 9 of 12 unsuccessful cases.

Analysis on the mistaken classification identified trends for improvements. In the case of false negatives – candidates with successfully filled positions being rated as not recommended – the unconventional career path was the primary reason for undervalued potential by the model. This was the case for candidates with unorthodox employment histories prior to law, for whom successfully employed trends were unrelated to applicable credential qualifications. On the other hand, false positives – unsuccessfully hired candidates with a recommended rating – had exemplary credentials and qualifications but were a cultural mismatch that was manifested after employment. These findings suggested that the model's limitations were primarily in capturing soft fit factors rather than credential assessment.

Phase 3 of user acceptance testing was conducted with the participation of five in-house legal managers. The managers carried out ten CVs with the tool and made judgments on the quality and utility of the resulting evaluation as well as its integration with current workflow. The feedback from the users pointed to several areas for improvement: managers wanted the ability to tailor the weighting for particular positions (for instance, emphasis on intellectual property experience for IP counsel positions); the text evaluation was at times too generic; and the tool was unable to discern information from atypical CV layouts.

Final touches incorporated feedback from users while being mindful of model integrity. A supportive prompt was created for users to articulate position-sensitive needs that influence the application of emphasis without changing the actual weighting. The textual production tasks were revised to require specific mentions within CVs and refrain from general comments. Formatting support was upgraded with more focused cues for extracting information from popular non-standard formatting such as LinkedIn PDF conversion.

The pilot testing stage lasted for around six weeks and included four major iterations on the instructional framework. The screening instrument achieved proper levels of accuracy with regard to classification at around 72–74% and showed high levels of agreement with the model (r = 0.91). Satisfaction with use was at satisfactory levels for acceptability criteria at a mean score of 4.1 out of a possible score of 5.0. Once the pilot testing was complete, the instrument was ready for use and protocols were laid out for monitoring and periodic model updates based on accumulated outcome findings.

# CHAPTER 5. ANALYSIS AND RESULTS

## 5.1 Descriptive Analysis of the Dataset

The overall sample consists of 269 legal professionals: 178 from the core TechCo sample and a further 91 from the LinkedIn group. This subchapter defines the population on important dimensions and thus serves to introduce the first systemic characterization of legal talent available within the IT industry in Ukraine.

Table 6 details the frequency of educational features found within the data set. A substantial number of candidates (67%) hold degrees from top law faculties within Ukraine, again reflecting the competitive nature of legal recruitment within the information technology industry. However, international educational attainment is a highly distinctive feature with merely 14% holding any form of educational award from outside Ukraine. This finding has important ramifications for the concept of international education signaling, with such characteristics being highly distinctive and thus having strong predictive power as a signal.

**Table 6**. *Educational Characteristics of Dataset (N = 269)*

| **Characteristic** | **n** | **%** |
|---|---|---|
| **University Tier** | | |
| Tier 1 (Top 5 Ukrainian) | 180 | 67% |
| Tier 2 (Other Ukrainian) | 51 | 19% |
| International | 38 | 14% |
| **Degree Level** | | |
| Bachelor only | 27 | 10% |
| Specialist/Master | 215 | 80% |
| LLM/PhD | 27 | 10% |
| **International Education** | | |
| Yes | 38 | 14% |
| No | 231 | 86% |

Characteristics of the experience imply that the workforce has a moderate tenured background within the legal field. Table 7 below highlights the experience profile for this dataset. The mean legal experience is calculated at 6.8 years (SD = 4.2), ranging from under one year to a total of 23 years. Prior IT company experience (being the strongest predictor within this regression model) was held by 34% of the recruitment candidates, establishing that a significant proportion of candidates have prior experience within the IT industry. Experience with a law firm was higher at 58%, consistent with established industry-standard career paths within the legal profession.

**Table 7**. *Experience Characteristics of Dataset (N = 269)*

| **Characteristic** | **Value** | |
|---|---|---|
| **Continuous Variables** | **M (SD)** | **Range** |
| Total Years Legal Experience | 6.8 (4.2) | 0.5 - 23 |
| Number of Employers | 3.4 (1.8) | 1 - 11 |
| Average Tenure (months) | 26.4 (14.7) | 6 - 84 |
| **Binary Variables** | **n** | **%** |
| IT Company Experience | 91 | 34% |
| Law Firm Experience | 156 | 58% |
| Big 4/International Firm | 67 | 25% |

| Characteristic | Value | |
|---|---|---|
| International Transactions | 118 | 44% |

English skills distribution highlights the importance of English ability in this set of candidates. A higher ability (levels 4-5) was found for 47% of the candidates, while Upper-Intermediate (level 3) was found for 31% and Intermediate (levels 1-2) and lower for 22% of the candidates. The bimodal distribution here with peaks for high and moderate ability can be given as one reason for the English Level being such a strong predictor: candidates tend to be either "internationally capable" or "domestically focused."

The qualitative variables derived from the result of narrative coding showed considerable variation within the sample. Table 8 shows the frequency distribution of qualitative indicators. The presence of Tech Industry Interest was strong (with a score of 2) for just 23% candidates, while for 41% candidates this quality was weak, and for another 36% this quality was absent altogether. It is observed that this particular quality is yet to be widely manifested as a strength in legal CVs, including in candidates applying for IT companies.

**Table 8**. *Distribution of Qualitative Variables (N = 269)*

| Variable | 0 (Absent) | 1 (Weak) | 2 (Strong) | Mean |
|---|---|---|---|---|
| Tech Industry Interest | 36% | 41% | 23% | 0.87 |
| Business Orientation | 28% | 44% | 28% | 1.00 |
| Growth Mindset | 31% | 47% | 22% | 0.91 |
| Adaptability | 33% | 42% | 25% | 0.92 |
| Proactivity | 39% | 43% | 18% | 0.79 |
| International Exposure | 42% | 35% | 23% | 0.81 |

*Note. Values show percentage of candidates at each coding level. Mean calculated on 0-2 scale.*

A comparison between the successful and unsuccessful candidates within the primary dataset shows stark differences. Table 9 highlights the mean differences between the two groups on the key parameters. Successful candidates score higher on English Level (4.2 compared to 3.1, $t = 5.84$, $p < .001$), IT Company Experience (56% compared to 26%, $\chi^2 = 12.4$, $p < .001$), and Tech Industry Interest (1.4 compared to 0.7, $t = 4.92$, $p < .001$). The bivariate differences mirror the patterns later observed in the multivariate regression models and reinforces the finding that the predictor parameters successfully distinguish between the two categories. Bivariate analysis was conducted only using the primary TechCo data source (n=178) because the validation of outcomes for individuals in the secondary data source would rely only on unvalidated reports available via professional networking sources. This helps avoid bias from unvalidated outcome data in the secondary sample.

**Table 9**. *Comparison of Successful vs. Unsuccessful Candidates on Key Variables*

| Variable | Successful (n=48) | Unsuccessful (n=130) | Test | p |
|---|---|---|---|---|
| English Level (M) | 4.2 | 3.1 | $t = 5.84$ | <.001*** |
| IT Company Exp (%) | 56% | 26% | $\chi^2 = 12.4$ | <.001*** |
| Int'l Transactions (%) | 67% | 38% | $\chi^2 = 10.8$ | .001** |
| Tech Interest (M) | 1.4 | 0.7 | $t = 4.92$ | <.001*** |
| Business Orient. (M) | 1.5 | 0.9 | $t = 4.21$ | <.001*** |
| Adaptability (M) | 1.4 | 0.8 | $t = 3.87$ | <.001*** |
| Avg Tenure (M months) | 31.2 | 23.8 | $t = 2.94$ | .004** |
| Law Firm Exp (%) | 54% | 59% | $\chi^2 = 0.34$ | .56 |

*Note. Primary dataset only (n = 178). **p < .01, ***p < .001*

### 5.2 Key Predictive Indicators Identified

In Chapter 4, the logistic regression analysis revealed that nine variables possess significant predictive power for success. The current subchapter explores further these predictors and explains how these predictors operate and the implications for screening.

IT Company Experience was found to be the best predictor (OR = 4.14, p = .003), with having more than four times the chance of success when this predictor is included. This finding offers strong support for Human Capital Theory: particular skills built within the tech industry can be transferred intact to the TechCo context. Attorneys with IT experience understand the pace, the relative importance, and the ways of communicating typical for tech companies. These lawyers are aware of the range of legal matters that occur (data protection and privacy, intellectual property, international contracts and deals involving venture capital investments).

The impact of experience with IT companies was more apparent when concurrent with a high level of English proficiency. Decision tree analysis showed that simultaneous possession of these two characteristics resulted in an 84% success rate compared with 52% for candidates with high English proficiency and no experience with IT companies and 19% for candidates lacking both characteristics. This interaction effect between the two variables identifies a complementary effect between them wherein IT experience provides industry-specific skills and English proficiency allows for efficient application of such skills at TechCo across international operations. From the screening process literature, simultaneous possession of both characteristics is a strong positive indicator.

English language proficiency (OR = 2.44, p = .004) was found to be the second strongest predictor among the quantitative aspects. Odds of success increased by a factor of about 2.4 for each level on the five-point scale. This is consistent with TechCo's international focus – products for international markets and business networks across several countries require English language proficiency for communicating with foreign business partners, preparation of contracts in English language, and functioning within English language regulatory frameworks. The decision tree cut-off at Level 4 (Advanced) implies the minimum required proficiency for this context.

International Transactions Experience (OR = 2.56, p = .016) was found to be highly predictive and significant after controlling for English proficiency. It appears that the candidates who have had exposure to international transactions are qualified to perform international transactions for TechCo without the necessity of direct supervision and with less likelihood of errors. Such candidates possess skills and knowledge that extend beyond language proficiency to legal system exposure and familiarity with international business practices and complex negotiations involving multiple parties.

In terms of the qualitative variables, Tech Industry Interest (OR = 2.14, p = .009) was found to be the best predictor for the narratives. If candidates can show evidence of actual interest in the tech industry, through mentions of products, trends within the tech industry, and tech industry-related hobbies, then they can increase their chance of success. This finding supports Person-Environment Fit Theory, whereby lawyers with actual interest in the tech industry tend to be more engaged with their work and tend to keep up with tech industry trends and can communicate well with members of other professions with similar industry interests.

Business Orientation (OR = 1.97, p = .012) was a significant predictor for success beyond legal competence. A resume emphasizing business and strategic thinking skills produced greater performance than one focusing on legal detail. This can be attributed to the specificity of the role of the general counsel. In contrast to lawyers at law firms, for whom billing for legal knowhow is sufficient, general counsel must be aware of the ways legal advice can enable business goals. Legal risk and business opportunities must be balanced by the ability to describe one's work among the candidates.

Adaptability (OR = 1.84, p = .015) captured the ability to deal with ambiguity and change. This predictor included indicators for role transitions, industry changes, and other types of responsibilities taken on. Given the quick pace at which tech firms grow and evolve and the associated shifts within legal priorities and needs, attorneys with a proven ability to be adaptable throughout their career and successfully complete transitions from one type of endeavor to another seem well-prepared to meet changing needs.

Average tenure showed a small but still significant effect (OR = 1.03 per month, p = .012). Individuals with more experience at previous companies were more likely to be selected by TechCo. For every increase of 12 months in average tenure, the chances of being selected were estimated to be approximately 43% higher. This finding can be utilized to validate the inclusion of stability at a job as one predictor for managing retention. However, this factor should be given less prominence than other criteria like IT experience and English skills.

Two factors came close to the threshold for conventionally being significant and were therefore included in the model: University Tier (OR = 1.68, p = .063) and Growth Mindset (OR = 1.63, p = .059). Both factors were included in the model for both their theoretical relevance and their role in fitting the model well. University Tier is founded on Signaling Theory, with graduation from a prestigious law school being indicative of cognitive and conscientious qualities necessary for good performance. Growth Mindset, implied by emphasis on learning and taking on a challenge, suggests a character predisposed to improvement for optimal outcomes.

Specifically, Law Firm Experience was nonpredictive (p = .42) after controlling for other predictors. The finding contradicts common wisdom governing recruitment for lawyers that considers law firms and other similar firms superior to other types of firms. It may be assumed that merely having law firm experience is irrelevant for candidates wanting to work IT lawyers; instead, the significance would be in the type of activities carried out (such as international contracts and IT law cases). Another reason is that being a Ukrainian advocate (Bar Admission) was nonpredictive (p = .67); this might be explained by the fact that IT lawyers do not frequently represent firms at court.

### 5.3 Model Performance Metrics

Internal validation was carried out on the primary TechCo sample (n = 178) via LOOCV. In this process, the model was trained on 177 observations and tested on the remaining one. This cycle was repeated for all observations. Table 10 below highlights the classification accuracy at the recommended threshold of 55 points.

**Table 10**. *Model Performance Metrics: Internal Cross-Validation (N = 178)*

| Metric | Value | 95% CI |
|---|---:|---:|
| Overall Accuracy | 74.2% | [67.1%, 80.4%] |
| Sensitivity (True Positive Rate) | 70.8% | [55.9%, 83.0%] |
| Specificity (True Negative Rate) | 75.4% | [67.0%, 82.6%] |
| Positive Predictive Value | 51.5% | [38.9%, 64.0%] |
| Negative Predictive Value | 87.5% | [79.6%, 93.2%] |
| Area Under ROC Curve (AUC) | 0.78 | [0.70, 0.85] |
| Cohen's d (Effect Size) | 1.24 | [0.89, 1.59] |

*Note. Classification threshold = 55. CI = Confidence Interval computed via bootstrap resampling (1000 iterations).*

The model has achieved AUC of 0.78 (95% CI: 0.70-0.85), reflecting strong discrimination power. In agreement with benchmarks suggested by Schmidt & Hunter (1998) and later corrected by Sackett et al. (2022), this outperforms the predictive validity for unstructured interviews (r ~ 0.38, AUC ~ 0.65; McDaniel et al., 1994) and comes close to the validity for the other method (r ~0.51, AUC ~0.71-0.76). In particular, this is accomplished on the sole basis of the CV data alone and with no additional information available from the interview.

The negative predictive value (NPV) is highly significant at 87.5%: If the model predicts not to take the candidate (score below 55), then this subject has a merely 12.5% chance to lead to a successful recruitment. The high negative predictive value supports the application of this model as a filter: the candidates with a score below the threshold can be dismissed with a relatively low risk that successful candidates will be missed.

External validation was carried out on a holdout sample (n = 25 members from the LinkedIn group with known outcomes) to evaluate the ability to generalize to candidates from outside TechCo. It produced 72% accuracy and AUC = 0.73 on this test set. While slightly subpar compared to the results from the previous section, these external outcomes lie well within the confidence interval for the internal estimates and can thus be said to be within acceptable bounds of generalizability. It is expected that there is at least some loss of ability compared to the previous validation, given the diversity in the legal community and may be represented by varying definitions for success.

The GPT model ran well in terms of agreement with the statistical model. With a sample size of 30 CVs assessed by both models, the correlation between the GPT model and the statistical model was $r = 0.91$ ($p < .001$), with a mean absolute deviation of 2.1. A high degree of agreement between the two models means that the design of the prompt successfully translated the regression formula and indicates that the GPT model closely replicates the statistical model's scores in this sample.

### 5.4 Comparative Analysis: Model Predictions vs. Real Recruitment Outcomes

In addition to overall model performance metrics, analysis on the model's performance on specific cases provides clarity on model strengths and weaknesses. In this subchapter, a detailed analysis of the confusion matrix is performed and specific trends within proper and improper classification are assessed.

Table 11 presents the confusion matrix from internal cross-validation. Of the 48 successful candidates, the model classified 34 (71%) successfully and incorrectly screened out 14 (29%). Of the other 130 candidates with unsuccessful attempts, the model screened out 98 (75%) successfully and incorrectly recommended 32 (25%). Overall, the model has a balanced approach between being sensitive and specific and has not optimized for one criterion at the expense of the other.

**Table 11**. *Confusion Matrix: Model Predictions vs. Actual Outcomes*

| | **Actual: Successful** | **Actual: Unsuccessful** | **Total** |
|---|---|---|---|
| **Predicted: Recommend** | 34 (True Positive) | 32 (False Positive) | 66 |
| **Predicted: Not Recommend** | 14 (False Negative) | 98 (True Negative) | 112 |
| **Total** | 48 | 130 | 178 |

*Note. Classification threshold = 55 points.*

Assessment on the false negative instances (14 out of the candidates being successfully hired while being unrecognized by the model) showed two main categories. In eight instances, candidates from unorthodox backgrounds were missing the marker for the credential from which the model based its prediction; candidates were highly successful due to exceptionally strong interview skills and a unique personal attribute that was unquantifiable from the resume. In six instances, candidates' strong traits were discovered on the job and would never emerge from the resume. It is clear from the above results that the model is weak at picking the soft aspects available only through direct human interaction.

On analysis of the false positives (32 candidates with no actual chance of being hired being suggested incorrectly), it was found that a majority were strong candidates with potential for cultural misfit and poor performance issues – parameters that cannot be judged from CV. Among these candidates, 17 seemed exemplary on paper with strong English Level, Information Technology Experience, and good qualitative parameters but were unable to meet the specific needs and parameters of the TechCo model. On the other end were 15 candidates with qualifications marginally above the minimum threshold and incapable of providing the skills and parameters that such qualifications represent.

The implication here for the confusion matrix is that the model should be used as a screening filter and NOT as a strict criterion for selection. The negative predictive validity is very high at 87.5%, and this is sufficient evidence for rejecting candidates with a score below 55. However, the positive predictive validity is relatively lower at 51.5% and this suggests that candidates with a score above 55 should be subjected to further screening since half of them may finally fail.

A threshold analysis was carried out to determine the impact of various decision thresholds on model performances. When the threshold is lowered to 45, more sensitivity is achieved at 85% (only misses 7 out of the successful candidates) at the expense of specificity at 58% (erroneously selects 55 unsuccessful candidates). On the other hand, a higher threshold at 65 raises specificity at 89% while reducing the model's sensitivity to 48% (misses half of all the successful candidates). The current threshold at 55 strikes a balanced approach between the TechCo recruitment context considerations for both false positives and false negatives.

### 5.5 Estimation of Time and Cost Savings

The usefulness of the screening model relies on its potential for generating significant efficiencies within the recruitment function. In this section, specific time and cost savings that can be realized with the screening approach would be assessed on the back of the recruitment metrics explained above and recruitment data from Chapter 1.

The baseline scenario, rooted within my personal experience as the CLO at TechCo, involves around 73 interviews every year with an average time of 134 minutes per interview (taking into consideration preparation and follow-up sessions). This amounts to around 163 hours every year for interviews alone. Out of these selected interviews, only 15 end with a successful recruitment process (a conversion rate of 20.5%), which means around 129 hours (a rate of 79.5%) relate to candidates that are either not hired or prove unsuccessful.

The screening model comes into play at the stage prior to conducting interviews with me after the review of CVs. Table 12 above highlights the expected effect based on scenarios for adoption. It is assumed in the analysis that candidates with scores below 55 are automatically excluded from being interviewed while candidates with higher scores proceed to the normal interview process.

**Table 12**. *Projected Time Savings from Screening Model Implementation*

| Metric | Baseline | With Model | Change |
|---|---|---|---|
| Candidates reviewed (CVs) | ~250 | ~250 | — |
| Candidates advancing to interview | 73 | 47 | -36% |
| Interview hours (annual) | 163 | 105 | -36% |
| Successful hires (annual) | 15 | 14 | -7% |
| Hours per successful hire | 10.9 | 7.5 | -31% |
| Conversion rate (interview to hire) | 20.5% | 29.8% | +45% |
| **Time saved (hours/year)** | — | **58** | |

*Note. Projections assume model sensitivity of 71% and specificity of 75% applied to baseline candidate flow.*

From the analysis above, the number of interviews expected to happen annually would come down from 73 to around 47 candidates, thus implying a decline of 36%. This is owing to the exclusion of candidates with scores below 55, with whom around 75% would have resulted in a failure. As a consequence of the model having imperfections with regard to its sensitivity measure, around four candidates with potential for being selected as successful employees would be eliminated (29% of the 14 false negatives in a typical year) and consequently reducing the number of successful employees from 15 to around 14.

The overall impact consists of time saved amounting to 58 hours per year (163→105 hours), with the downside being the loss of one successful hire. Whether this is a good deal would depend on the relative cost of the time of a senior legal manager and the cost of the lost hire. At a senior legal manager hourly cost of $100 (including salary, benefits, and overhead), the time savings represent approximately $5,800 annually. If missed successful hire costs are less than this amount, the expected direct annual value of the model will be positive (not including the indirect value effect which may reach much higher amounts).

The improvement in conversion rate from 20.5% to 29.8%, which is a relative increase of 45%, means that there is a qualitative benefit over and above the time-saving process. Better conversion rates mean that interview time is being used more productively because more candidates interviewed end up being hired. This is a time-saving benefit that can lead to other benefits such as a reduced need for the candidate pipeline and an improvement in the candidate experience process because candidates reach a decision sooner.

Scaling these estimates with regard to the Ukrainian information technology market may lead to a potential impact multiplier effect. If fifty similar legal departments spend equal time on recruitment activities, then industry-wide interview time per year is close to 8,150 (163 × 50). A 36% reduction would lead to approximately 2,900 industry-wide time savings every year, which is close to releasing 1.5 full-time senior legal managers for actual legal work.

The cost of deployment for the model is minimal. In the case involving the GPT model, the tool requires just a subscription to the ChatGPT Plus (around US$20 per month) and minimal computation time for each resume (less than 30 seconds). It does not require any investment for the maintenance of infrastructure and human resources, as well as additional training. The biggest expense would be for the development process outlined within this research and can be shared over several firms adopting the process.

It should be pointed out that these estimates are based upon several assumptions that may not be fully met. The parameters for recruitment are based on TechCo, and this may not be the case for other companies. The model's performance on external benchmarks showed some loss, and this means that context-based adaptation may be necessary. In addition to this, these estimates are based upon the assumption that the screening system is being used for its intended purpose, and this is for filtering and not for decision-making on its own. In any case, the conclusion is supported that the screening model is highly useful with regard to improving the efficiency of recruitment.

Findings from this chapter show that the screening model produced by AI meets its main task: to find predictive patterns on CVs separating successful from unsuccessful legal professionals. The model with nine variables has a substantial discrimination power (AUC = 0.78), and its efficiency is either at the same or even higher levels than traditional methods used for screening. All of the above results confirm existing meta-analytical data showing that mechanical (statistic) integration of data leads to better performance compared to clinical methods in selection and admission decisions (Kuncel et al., 2013). Findings confirm hypotheses from Human Capital Theory, Signals Theory, and Person-Environment Fit Theory and imply at the same time that some categories with traditionally assumed high merit (like law firms and bar pass status) lack predictive power for the Ukrainian IT market.

# CHAPTER 6. CONCLUSIONS

## 6.1 Summary of Key Findings

The research offers evidence that both supports and undermines general perceptions on legal recruitment. The core empirical finding is that CV screening is indeed capable of producing predictive validity for legal recruitment outcomes in the Ukrainian IT companies. More specifically, the model produced an AUC of 0.78, meaning that given one successful and one unsuccessful candidate, the model would successfully identify the former with a probability of about 78% on average. This is a desirable property compared to other methods available and well outperforms a randomly guessing model.

The predictive structure that was produced demonstrates a marked departure from the traditional yardstick used within legal recruitment. Industry experience, with specific attention given to previous employment within the IT industry, has been identified as the primary predictor within this model, with candidates having such experience being over four times more likely to succeed. The finding contradicts the traditional prestige ranking system, wherein candidates from more prestigious legal firms would be selected with higher precedence. English proficiency was found to be the second significant indicator, wherein proficiency is necessary and considered not merely as additional qualification criteria for international interaction. Both factors interact with each other in a complementary process, wherein candidates with both factors achieved well above the predicted rate for each factor alone.

Perhaps more enlightening than the correlate of success is what does not correlate with success. Experience at a law firm, conventionally perceived as constituting foundational preparation for legal practice, lacked any independent predictive power once other predictor variables were held constant. Also lacking independent predictive power was Ukrainian bar admission as a means distinguishing between those who were and were not successful. The implication here is that screening processes focusing on these more traditional parameters amount to filtering noise and not signal.

In addition to these findings, the research found that qualitative aspects of CVs, namely the ways that candidates perceive their work experiences relative to objective employment history, add predictive validity. Variables such as actual interest in technology on the part of the candidate, the business application of legal expertise, and personal indicators of resilience were all found to add predictive validity beyond objective qualifications. The implication is that there is more information on the CV than is being tapped by traditional screening processes and that self-presentation provides insight into person-organization fit that is not gathered from qualifications.

The application of these outcomes as a GPT tool for screening showed empirical validity. It processes CVs efficiently and achieves validation rates suitable for a filter-stage screening tool. The estimated efficiency gain per year at 58 hours for legal managers and minimal risk of missing successful candidates confirm the economic viability for adoption. However, with a high negative predictive value of 87.5% and moderate positive predictive value at 51.5% for the screening tool, exclusion and inclusion criteria eliminate and select candidates fitting for advancement while allowing for managerial discretion during the final screening and recruitment stages.

## 6.2 Contribution to Theory and Practice

This research offers improvements on multiple fronts: (1) theoretic knowledge with regard to the recruitment process, (2) methodology for applying AI-based recruitment solutions, and (3) tool-based application for managing legal talent at tech companies operating globally from emerging markets.

Theoretically, this work offers a more detailed understanding of the processes under which Human Capital and Person-Environment Fit Theories interact within signaling to offer a predictive framework for professionals. It is clear from this study that general human capital (broad legal knowledge and skills) is outweighed by specific human capital (industry experience) as predictors for professionals and thus offers a more refined model within Becker's framework with more general application than legal recruitment. In terms of Signaling Theory, the evidence suggests a more complex application than traditional notions allow, demonstrating that traditional legal indicators (prestigious law firms and bar admission) lack predictive validity when the underlying skills they purport to signal are irrelevant to the role requirements. On the other hand, the candidates' narrative self-presentation demonstrated to be predictive, pointing to the existence of subtle, implicit signaling channels beyond formal credentials. Person-environment fit offers strong predictive evidence within this study and for the explanation for false positives for qualified candidates.

The linking of the three theory approaches to a combined prediction model is a conceptual contribution in its own right. None of the theory approaches alone is a complete predictor for hiring success: credentials provide a basis for qualifications, experience offers capability and fit will determine if qualified and capable people will prosper in the particular environment represented by the organization in question. The model fits into the developing theory in the area of personnel selection, where a combination of methods is sought instead of a single predictor.

Methodologically speaking, the study proves to be a working model for creating AI-based screening tools in a data-scarce professional environment. The combination of traditional statistical inference to find the important predictive variables and the application of an LLM to implement those variables works as a model that can be replicated in similar domains. The technique allows for model explainability (the users are able to understand why certain grades will be given to certain applicants) and the application of natural language processing (the tool finds values in unstructured CVs without constant manual feature extraction). The testing model, involving cross-validation in the sample and test sample validation, gives due credibility to the tool in terms of its merit.

The practical contribution is the screening instrument and the accompanying framework for implementation. For legal managers working in the Ukrainian IT sector, the study provides an instantly applicable technology to improve the efficiency of hiring. Additionally, aside from the screener itself, the results offer insights into recruitment practices that can be applied regardless of whether the processes of the given organization entail the use of a GPT screening tool.

The current study offers the first systematic characterization of legal talent in the Ukrainian IT sector. The results of the descriptive review of 269 legal professionals provide the baseline information necessary to characterize the group members in terms of their educational attainment, experience, language skills, work paths and trajectories. This work yields a description that is broadly applicable and goes well beyond its immediate use in predictive modeling.

Finally, this project contributes to the discussion regarding the responsible application of AI to high-stakes human decision-making processes. The resulting framework regarding the issues of consent, data privacy, mitigating biases, and the need for transparency offers a template for how to responsibly implement AI recruitment software in a company. The careful caveats and the decision to characterize the tool as a decision aide adopt an appropriately measured and cautious approach of the application of AI to enhance human decision-making in employment.

### 6.3 Limitations of the Study

Intellectual honesty requires a frank assessment of the study's limitations, and these tend to weaken the confidence that should be vested in the results and circumscribe the domain in which the results should be applied.

The most obvious limitation pertains to the sample size and number of observations. With only 48 success observations out of a total of 178 in the basic dataset, the statistical power to distinguish smaller effects is suspect and the sample stability of the coefficient estimates unclear. The model designed to predict success employs nine predictors and approaches the limits of its sample size. Thus, certain factors known to be predictive of success are left out because the statistical power to distinguish their effect is lacking, and the estimates of the predominant factors may be unstable and misleading to a certain degree. Replication with larger samples is necessary before the specific predictive weights can be considered definitive.

Single-source development leads to issues of generalizability. The model was developed based upon hiring success at TechCo, and TechCo may be unlike other Ukrainian firms in the IT sector in terms of their culture, the importance of the legal roles, the criteria for hiring success, and a variety of other factors that in turn affect success determinants. While the model has been validated through the sample from LinkedIn group, the primary dataset is much more detailed than the secondary sample.

The outcome measure has its shortcomings too. The basis of success in terms of retention in the organization at the end of two years and performing well is reasonable but has its flaws. The success criterion doesn't consider strongly performing employees with less experience who quit for better prospects but includes employees with longer tenure and average performers who deliberately avoided getting terminated instead of working well in the company. Other methods of defining success might generate different results. Moreover, the success criterion is applicable only to the shortlisted candidates, and the model doesn't validate the predictions made for the rejected candidates and how they might perform successfully given the chance.

The coupling of researcher involvement in hiring decisions and the construction of predictive tools leads to the possible introduction of biases. My capacity as the CLO at TechCo means that the hiring decisions foretold by the tool will partially depend upon my assessment of the quality of the applicants. The possible presence of biased decision-making will thus lead the tool to impart and perpetuate the said biases instead of alleviating them. The use of objective performance measures combined with retention data partially mitigates this issue; however, the tendency for self-confirming bias cannot be entirely excluded.

The temporal scope generates an element of indeterminacy in terms of long-term validity. The data range is between 2017 and 2025, a time of great turbulence for the environment in Ukraine, which saw the beginning of a full-scale war in 2022. The variables that were performing well in the previous environment could potentially behave differently given the current state of affairs that have been shaped by the rising use of remote work and talent displacement. The model assumes an unchanging environment, but this assumption requires constant verification against a changing setting.

The results for the qualitative coding of narrative variables, while demonstrating an acceptable inter-rater reliability, add a degree of subjective variability to the predictors characterized as objective in nature. Different raters tend to attach a differing measure of their assessment of technical interest and business acumen to the same CV, and the GPT function is unlikely to strictly capture human assessments made in the coding process. The presence of a robust covariance relationship between GPT scores and model scores ($r = 0.91$) reveals a non-perfect level of congruence and the presence of systematic variations in the way the tool rates candidates against the calibrated model.

The GPT platform brings in a number of dependencies and risks. Changes in the model by Open AI might affect the robustness of the evaluated results, the platform might be unavailable and there might be changes in the costs associated with the platform, while the fact that GPT is proprietary software means that the internal processing operations in the platform cannot be deeply audited. Using the tool, the organizations involved incur the platform risks that fall out of their control.

The current study is not an investigation into the entire process of hiring. By focusing solely on the screening process through CVs, issues in the optimization of other processes in hiring, such as the design of interviews and the effectiveness of the onboarding process, are left unaddressed. A candidate may be able to pass through the screening process but fail in the interviews, or they may be hired in the organization but fail due to a lack of proper support, and this falls outside the realm of the screening process optimization model. The tool aims to optimize one element in a multi-step process and should in no way be considered a full solution to the hiring issues in the IT legal sector.

### 6.4 Implications for Future Research

The above limitations suggest promising avenues for future research, and the positive results worth replicating are encouraging.

The highest priority area for studies is multi-organizations replication. The studies focusing on legal recruitment in other Ukrainian IT companies will reveal whether the results from TechCo are transferable across organizations or if they describe organization-specific associations. The combination model, in which a number of firms provide their data while keeping it confidential, will provide enough samples to perform more precise statistical calculations. Moreover, the model will make it possible to check moderating phenomena and whether the size of a company, the level of global presence, and the scope of the legal department affect the importance of the screening variables.

Additional important area of work involves the longitudinal validation process because studies conducted across the lifetime of the applicants screened by the tool will provide a possible way to know the predictive validity free from the prospective biases known to affect the current validity studies. The research will be able to evaluate if the screening results predict success in the long run and if leadership will emerge based upon the results of the prediction processes involving success and long-term performance results.

The cross-functional extension is designed to test the applicability of the methodology to other specialized professions in the technology sector. Product managers, compliance officers, and high-level finance executives face similar screening issues, where the application of credentials-based hiring might be misaligned in terms of its reliance on predictor variables of success in the function in question. The application of the methodology to such functions will enhance its applicability and test the generic nature of the theoretical model to establish if the noted phenomena are function-specific in the legal domain or also in the broadly generic domain of the hiring of professionals in the technology sector.

There appears to be a need for a more detailed review of the results from the narrative predictors. The fact that candidates' narrative framing of their work experience is predictive of success independently of their level of observable performance suggests that narrative prediction instruments and cognitive/personality tests could use different mechanisms. This result deserves closer examination to understand the relationship between these instruments better and make narrative assessment instruments more reliable.

The null findings concerning the experience in a law firm and bar admission to practice the law should be investigated in a targeted manner. It is crucial to understand whether the current results are isolated to the Ukrainian IT sector or if they represent a more general pattern in the legal hiring process. Studies in other settings, ranging from other countries to other spheres of the legal profession, will shed light upon whether the traditional qualifications have lost their explanatory value in the general case, or if the phenomenon is limited to a current environment.

Studies regarding bias and fairness will take an ever more important significance in the advent of AI screening tools in wider applications. Future studies should examine the screening model to determine if whether proxy discrimination emerges through seemingly neutral predictors, in addition to testing if biased regression methods of mitigation will diminish the negative impacts while preserving the predictive validity. All the above analyses require the collection of demographic data that was not encompassed in the current work.

AI screening architectures might be compared and contrasted to guide the choice of technology tools. The current project explored a specific methodology – the application of statistical methods to find predictors and the implementation of the predictor in GPT-based architecture, while there are other possible alternatives too. The methods involving the training of fine-tuned language models trained on labeled examples of CVs and outcomes, ensembles based upon the combination of a variety of algorithms, and the combination of the structured and unstructured methods might serve as better alternatives.

Finally, future studies focusing on human-AI collaboration in the hiring process should aim to investigate issues that go well beyond the performance of the AI algorithm itself. For instance, how precisely do hiring managers make use of the output of screening tools in the hiring process? Do the recommendations of the tool guide managerial decision, compete with it, or are ignored altogether? Do screening tools affect the way the hiring managers make decisions and do they improve or worsen the quality of such decisions? It is a fact that the understanding of the total system – the combination of the tool and its application in its environment – is as critical as the tool's technical performance.

The capstone project springs from a real-world dilemma: too much interviewing of people unlikely to be hired combined with a lack of confidence in the current screening methods to properly find the right people. The AI screening tool designed in this project represents one solution to this dilemma and a collection of results that shed light upon why the traditional methods are less effective. The AI screening tool designed in this project is in no way a complete solution to the hiring dilemma but instead represents an important step towards a more informed hiring process in the legal sector. The importance of the screening process will and continues to be a critical area of focus for Ukrainian IT companies in their fight to compete in the global economy because the ability to find suitable legal professionals well-versed in cross-border complexity will be an important priority in the years to come.

## APPENDIX A. DATA COLLECTION INSTRUMENTS

This appendix contains the forms designed to collect information from the candidate CVs to be coded and analyzed. They contain the definitions and rules of the variables from Chapter 3 (Tables 1 and 2), and such forms are designed to be followed while working with the tables. The rules for the success classification are explained in subchapter 4.2.

In total, there are three instruments: the CV Coding Form (A.1), the Hiring Outcome Assessment Form (A.2), and the Inter-Rater Reliability Protocol (A.3).

### A.1 CV Coding Form

The form should be filled out for each candidate. Apply the rules for coding, according to Table 1 for quantitative variables and Table 2 for qualitative variables. The precise wording from the CV supporting the qualitative assessments should be recorded.

Candidate ID: ____________ Date: ____________ Coder: ____________

**Quantitative Variables** *(coding rules: Table 1, Chapter 3)*

| Variable | Code | Variable | Code |
|---|---|---|---|
| University Tier | ____ | English Level | ____ |
| Degree Level | ____ | English Certification | ____ |
| International Education | ____ | Additional Languages | ____ |
| Total Years Legal Experience | ____ | Bar Admission | ____ |
| IT Company Experience (years) | ____ | International Qualification | ____ |
| Law Firm Experience (years) | ____ | Specialized Certifications | ____ |
| Big 4 / International Firm | ____ | IP/IT Practice Experience | ____ |
| Number of Employers | ____ | Corporate/M&A Experience | ____ |
| Average Tenure (months) | ____ | International Transactions | ____ |

**Qualitative Variables** *(coding indicators: Table 2, Chapter 3)*

| Variable | 0-2 | Supporting Evidence from CV |
|---|---|---|
| Tech Industry Interest | | |
| Business Orientation | | |
| Growth Mindset | | |
| Adaptability | | |
| Proactivity | | |
| Team Collaboration | | |
| Quantified Results | | |
| Leadership Evidence | | |
| Problem-Solving Examples | | |
| International Exposure | | |

| | | |
|---|---|---|
| Entrepreneurial Signals | | |
| Values Alignment | | |

**Notes/Ambiguities:** ______________________________________________

### A.2 Hiring Outcome Assessment Form

The assessment should be made based upon HR records and performance results. Subject to the success definition, a candidate will be regarded as a success if they have been made an offer AND have reached the necessary term of at least 24 months of tenure (or current employment) AND performance grades of 3 ("Meets Expectations") or higher.

Candidate ID: ___________ Assessment Date: ___________

| **APPLICATION & EMPLOYMENT DATA** |
|---|
| Application Date: __________ Position Applied: _______________________ |
| Offer Extended: ☐ Yes ☐ No Offer Accepted: ☐ Yes ☐ No ☐ N/A |
| Start Date: __________ End Date (if departed): __________ |
| Current Status: ☐ Currently Employed ☐ Voluntary Departure ☐ Involuntary |
| Total Tenure: ________ months |

| **PERFORMANCE DATA (hired candidates only)** |
|---|
| Most Recent Rating: ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 Rating Period: __________ |
| Promoted During Tenure: ☐ Yes ☐ No ☐ N/A |

| **SUCCESS CLASSIFICATION (per Subchapter 4.2)** |
|---|
| ☐ SUCCESSFUL: Offer + (tenure ≥24 mo OR current) + performance ≥3 |
| ☐ SUCCESSFUL: Offer extended, declined by candidate |
| ☐ UNSUCCESSFUL: Hired but tenure <24 mo OR performance <3 |
| ☐ UNSUCCESSFUL: Not offered position |

**FINAL CLASSIFICATION:** ☐ SUCCESSFUL ☐ UNSUCCESSFUL

### A.3 Inter-Rater Reliability Protocol

This protocol specifies the consistency of programming for qualitative variables. As discussed in subchapter 3.3.2, inter-rater reliability was assessed in a sample of 20 random CVs, and the value of Cohen's kappa exceeded 0.7 for each variable.

***Procedure***

Twenty CVs were purposively sampled from the primary dataset and stratified based upon outcome (10 success and 10 failure) and coded independently according to the definitions outlined in Table 2 by two independent coders without prior knowledge of candidate outcomes. The value of Cohen's kappa was determined for each variable. Values of Kappa less than 0.7 required a re-definition of the coding definitions and re-testing until the value of Kappa reached the target value of 0.7. Coding disagreements in the principal dataset were resolved through discussion to achieve a consensus.

***Reliability Results***

| Variable | Cohen's κ | Interpretation |
|---|---|---|
| Tech Industry Interest | 0.78 | Substantial |
| Business Orientation | 0.74 | Substantial |
| Growth Mindset | 0.71 | Substantial |
| Adaptability | 0.76 | Substantial |
| Proactivity | 0.72 | Substantial |
| Team Collaboration | 0.70 | Substantial |
| Quantified Results | 0.85 | Almost perfect |
| Leadership Evidence | 0.77 | Substantial |
| Problem-Solving Examples | 0.73 | Substantial |
| International Exposure | 0.82 | Almost perfect |
| Entrepreneurial Signals | 0.79 | Substantial |
| Values Alignment | 0.70 | Substantial |

*Note. Interpretation per Landis & Koch (1977): 0.61-0.80 = substantial; 0.81-1.00 = almost perfect.*

***Discrepancy Resolution Notes***

The discrepancies with the highest incidence involved the categories of Values Alignment and Team Collaboration, where the coders differed in their assessment of whether generic terms of a professional nature (such as a "team player") should be scored 1 or 0. Resolution: generic terms that lacked specific examples were assigned a score of 0, while examples of cooperative activity were assigned a score of 1 to 2 based upon the level of importance emphasized.

For the category of Tech Industry Interest, discrepancies occurred where the job applicants listed their work for technology clients but failed to indicate their personal interest in the sector. Resolution: simply working for clients was a score of 0 unless coupled with an expression of genuine interest beyond the level of a professional duty to do so.

The category of Quantified Results evidenced the highest levels of reliability with a value of $\kappa = 0.85$ due to the clear presence or absence of numerical evidence.

## APPENDIX B. GPT SYSTEM PROMPT

This appendix lists the full system prompt used to customize the GPT-5.2-based screening tool. The full system prompt has a total of ~3,500 tokens and is divided into six functional sections, as described in subchapter 4.4.2 above. The full system prompt listed in the following paragraphs is a direct copy of the one implemented in the system, with slight changes made solely to improve its readability. Those interested in replicating the screening tool should employ the identical prompt structure with GPT-5.2 or comparable LLMs.

**— BEGIN SYSTEM PROMPT —**

### Section 1: Role Definition

You are a screening assistant expert in the legal recruitment sector, specializing in screening CV/resume data to provide structured assessments to predict the success of candidates applying for in-house counsel recruitment in the technology sector in Ukraine. The task you need to focus on is the CV/resume screening and provide assessments based on empirical and validated recruitment criteria assembled from the testing of 269 legal professionals in the legal recruitment sector.

The assessments should serve to assist human decision-making and should in no way be a replacement. No recommendations should be considered in employment decisions unless vetted by human hiring professionals.

Context: The assessment is related to legal positions in a Ukrainian IT company that has global presence. The success factor in such legal positions requires legal expertise in the domain and a business acumen that matches the fast-paced environment of the technology sector.

### Section 2: Information Extraction Rules

When analyzing a CV, extract and categorize the following information:

QUANTITATIVE VARIABLES:

- University_Tier: Classify as 1 (Top 5 Ukrainian), 2 (Tier 2), or 3 (Other Top 5: Taras Shevchenko, Yaroslav Mudryi, Ivan Franko, Kyiv-Mohyla, KNEU
- Degree_Level: 1 (Bachelor), 2 (Specialist/Master), 3 (PhD/Candidate)
- International_Education: 0 (No) or 1 (Yes - LLM, exchange, foreign degree)
- Total_Legal_Experience: Years since first legal position (numeric)
- IT_Company_Experience: Years at technology/IT companies (numeric)
- Law_Firm_Experience: Years at law firms (numeric)
- Big4_International: 0 (No) or 1 (Yes - Big 4 or international firm)
- Number_Employers: Count of distinct employers
- Average_Tenure: Mean months per employer
- English_Level: 1 (Basic), 2 (Intermediate), 3 (Upper-Int), 4 (Advanced/Native)
- English_Certification: 0 (None stated) or 1 (IELTS/TOEFL/Cambridge/etc.)
- Additional_Languages: Count of languages beyond Ukrainian/Russian/English
- Bar_Admission: 0 (No) or 1 (Ukrainian advocate license)
- International_Qualification: 0 (No) or 1 (Foreign bar/qualification)
- Specialized_Certifications: Count (CIPP, CIPM, PMP, tech certs, etc.)
- IP_IT_Experience: 0 (No) or 1 (IP, IT law, data protection experience)
- Corporate_MA_Experience: 0 (No) or 1 (M&A, corporate transactions)
- International_Transactions: 0 (No) or 1 (Cross-border deal experience)

QUALITATIVE VARIABLES (score 0-2 based on CV evidence):

- Tech_Industry_Interest: Genuine enthusiasm for technology sector
  0=Absent, 1=Mentioned tech clients, 2=Active engagement/passion evident
- Business_Orientation: Understanding of commercial/business context
  0=Absent, 1=Some commercial awareness, 2=Strong business focus
- Growth_Mindset: Evidence of continuous learning and development
  0=Absent, 1=Some professional development, 2=Strong learning pattern
- Adaptability: Capacity for change and new environments

```
     0=Absent, 1=Some transitions, 2=Clear adaptability evidence
   - Proactivity: Initiative beyond assigned responsibilities
     0=Absent, 1=Some initiative, 2=Strong proactive pattern
   - Team_Collaboration: Evidence of collaborative work style
     0=Absent, 1=Team mentioned, 2=Specific collaboration examples
   - Quantified_Results: Numerical achievements and metrics
     0=Absent, 1=Some numbers, 2=Multiple quantified accomplishments
   - International_Exposure: Cross-border work or multicultural experience
     0=Absent, 1=Some international work, 2=Substantial global experience
   - Entrepreneurial_Signals: Startup involvement, side projects, innovation
     0=Absent, 1=Minor signals, 2=Clear entrepreneurial activity
```

### Section 3: Scoring Algorithm

```
The total score will be measured in a 100-point scale based on the weights derived for
logistic regression analysis:

QUANTITATIVE COMPONENT (55 points maximum):
   - IT_Company_Experience: 0-15 points
     (0 yrs=0, 1-2 yrs=5, 3-4 yrs=10, 5+ yrs=15)
   - Big4_International: 0 or 10 points
     (Yes=10, No=0)
   - International_Education: 0 or 8 points
     (Yes=8, No=0)
   - English_Level: 0-8 points
     (Level 1=0, Level 2=2, Level 3=5, Level 4=8)
   - IP_IT_Experience: 0 or 7 points
     (Yes=7, No=0)
   - International_Transactions: 0 or 7 points
     (Yes=7, No=0)

QUALITATIVE COMPONENT (45 points maximum):
   - Tech_Industry_Interest: 0-8 points
     (Score 0=0, Score 1=4, Score 2=8)
   - Business_Orientation: 0-8 points
     (Score 0=0, Score 1=4, Score 2=8)
   - Growth_Mindset: 0-8 points
     (Score 0=0, Score 1=4, Score 2=8)
   - Adaptability: 0-7 points
     (Score 0=0, Score 1=3.5, Score 2=7)
   - Proactivity: 0-7 points
     (Score 0=0, Score 1=3.5, Score 2=7)
   - Quantified_Results: 0-7 points
     (Score 0=0, Score 1=3.5, Score 2=7)

RECOMMENDATION THRESHOLDS:
   - 70-100: STRONGLY RECOMMEND - High probability of success
   - 55-69:  RECOMMEND - Good potential, proceed to interview
   - 40-54:  CONSIDER WITH CAUTION - Some concerns, discuss with team
   - 0-39:   DO NOT RECOMMEND - Low predicted success probability
```

### Section 4: Qualitative Assessment Guidelines

```
When assigning qualitative scores, apply these interpretation rules:

CONSERVATIVE CODING PRINCIPLE:
```

When evidence is ambiguous, assign the lower score. Absence of evidence is not evidence of absence, but CVs are self-promotional documents; if a candidate does not highlight a quality, assume it is not a strength.

TECH INDUSTRY INTEREST INDICATORS:
- Score 2: Personal tech projects, tech blog/writing, startup involvement, expressed passion for technology, voluntary tech sector moves
- Score 1: Tech company clients, IT-related legal work, tech certifications
- Score 0: No technology-related content beyond generic practice areas

BUSINESS ORIENTATION INDICATORS:
- Score 2: Revenue/cost impact described, commercial strategy involvement, business development activities, P&L awareness
- Score 1: Contract value mentions, client relationship management
- Score 0: Purely legal/technical descriptions without business context

GROWTH MINDSET INDICATORS:
- Score 2: Multiple certifications over time, diverse role progression, learning activities mentioned, career pivots with skill acquisition
- Score 1: Some professional development, occasional training
- Score 0: Static skill set, no development activities mentioned

CULTURAL FIT WARNING SIGNS (note but do not penalize without evidence):
- Very short tenures (<12 months) at multiple employers
- Exclusively government/bureaucratic background
- No evidence of collaborative work in any role
- Rigid career progression with no lateral moves

## Section 5: Reference Profiles

For calibration, the anonymized versions of the reference profiles will be used. Each profile refers to an existing candidate in the training data.

=== HIGH SUCCESS PROBABILITY PROFILES (Scores 75-90) ===

PROFILE HS-1 (Score: 87, Outcome: Successful, 4+ years retained)
- Education: Top 5 university, Master's degree, LLM from UK university
- Experience: 6 years total, 4 years IT company, 2 years international firm
- Languages: English C1 certified (IELTS 7.5), German basic
- Practice: IP/IT law, data protection, international transactions
- Qualitative: Strong tech interest (personal blog), business metrics cited, continuous certifications (CIPP/E), proactive project leadership

PROFILE HS-2 (Score: 82, Outcome: Successful, promoted within 2 years)
- Education: Tier 2 university, Specialist degree
- Experience: 8 years total, 5 years IT companies, Big 4 background
- Languages: English C1 (no certification), Polish intermediate
- Practice: Corporate/M&A, venture capital, commercial contracts
- Qualitative: Business-focused descriptions, deal values quantified, cross-functional team leadership, startup advisory experience

PROFILE HS-3 (Score: 78, Outcome: Successful, currently employed 3+ years)
- Education: Top 5 university, Master's degree, exchange semester abroad
- Experience: 5 years total, 3 years IT company, 2 years boutique firm
- Languages: English B2-C1, French basic
- Practice: Employment law, corporate, compliance
- Qualitative: Adapted from litigation to in-house, built compliance program, quantified policy implementation results, team collaboration emphasized

=== MODERATE SUCCESS PROBABILITY PROFILES (Scores 55-74) ===

```
PROFILE MS-1 (Score: 68, Outcome: Successful, adequate performance)
   - Education: Tier 2 university, Specialist degree
   - Experience: 7 years total, 2 years IT company, 5 years law firm
   - Languages: English B2
   - Practice: General corporate, contracts
   - Qualitative: Some tech client work, limited business language,
     steady career progression, team player mentioned generically
PROFILE MS-2 (Score: 62, Outcome: Successful, retained but not promoted)
   - Education: Top 5 university, Bachelor's + Specialist
   - Experience: 4 years total, 1 year IT, 3 years mixed practice
   - Languages: English B2, pursuing C1
   - Practice: IP basics, general corporate
   - Qualitative: Interested in tech (mentioned in objective), growth evident
     through certifications in progress, adaptable but limited examples
PROFILE MS-3 (Score: 58, Outcome: Unsuccessful, departed at 18 months)
   - Education: Tier 2 university, Master's degree
   - Experience: 10 years total, 0 years IT, 8 years large law firm
   - Languages: English C1 certified
   - Practice: Litigation, dispute resolution, arbitration
   - Qualitative: Strong credentials but no tech affinity evident,
     business orientation limited, rigid career in litigation only

=== LOW SUCCESS PROBABILITY PROFILES (Scores below 55) ===
PROFILE LS-1 (Score: 45, Outcome: Unsuccessful, not hired)
   - Education: Other university, Bachelor's degree
   - Experience: 3 years total, 0 years IT, government + small firms
   - Languages: English B1
   - Practice: Administrative law, general practice
   - Qualitative: No tech indicators, no business language, limited growth
     evidence, purely compliance-focused descriptions
PROFILE LS-2 (Score: 38, Outcome: Unsuccessful, rejected at screening)
   - Education: Other university, Specialist degree
   - Experience: 6 years total, 0 years IT, 6 years government/state enterprise
   - Languages: English B1-B2 (self-assessed)
   - Practice: Regulatory compliance, licensing, permits
   - Qualitative: No tech, no business orientation, no adaptability signals,
     bureaucratic language throughout, no quantified achievements
PROFILE LS-3 (Score: 42, Outcome: Unsuccessful, terminated at 8 months)
   - Education: Tier 2 university, Master's degree
   - Experience: 5 years total, 1 year IT (departed quickly), job-hopping pattern
   - Languages: English B2
   - Practice: Various - litigation, corporate, employment, IP (surface level)
   - Qualitative: Breadth without depth, 5 employers in 5 years,
     no sustained achievements, no team collaboration evidence
```

## Section 6: Ethical Constraints

MANDATORY ETHICAL REQUIREMENTS:

1. NON-DISCRIMINATION

   Do not take into consideration and make any mention of gender, age, marital status, parental status, ethnicity, religious belief, disability, pregnancy, military service, and any other characteristic that is considered protected. Evaluate based solely on qualifications related to the job.

2. TRANSPARENCY

Justifications should be explicit for each score given. Each recommendation should be able to be traced to the content in the CV. Avoid making inferences regarding information that is unaccounted for in the CV.

3. HUMAN OVERSIGHT

Always make clear the purpose is to provide a tool to guide decision-making only. The hiring decision should be left to the qualified human assessors. The purpose of the process is to draw attention to the relevant information, and not to make the hiring decision.

4. UNCERTAINTY ACKNOWLEDGMENT

When ambiguity and incompleteness occur in the information, the uncertainty should be clearly acknowledged. Use confidence qualifiers "CV indicates…", "Evidence suggests…", and "Unable to determine from the available information…".

5. GDPR/DATA PROTECTION ALIGNMENT

Exclusively process the data in the CV and do not attempt to gather any additional information. Do not retain any information from any previous assessments. Every assessment should be performed and based solely upon the input received.

6. BIAS MONITORING

If there appear to be patterns in the results that might indicate a possible systematic bias (e.g., consistently lower achievement associated with certain levels of educational attainment that are not supported by the achievement results), such issues should be brought to the attention of a human reviewer.

PROHIBITED ACTIONS:

- Drawing conclusions regarding the success/failure chances of a candidate.
- Making a recommendation for rejection based solely on the quantitative score.
- Making inferences regarding the personality of an individual based upon the limited available information contained in a CV.
- Using informal language or humor in assessment design and communications.
- Giving candidates direct feedback.

### Output Format Specification

For each CV assessment, produce output in this exact format:

```
====================================================================
CANDIDATE SCREENING ASSESSMENT
====================================================================
OVERALL SCORE: [XX]/100
RECOMMENDATION: [STRONGLY RECOMMEND / RECOMMEND / CONSIDER WITH CAUTION /
DO NOT RECOMMEND]
--------------------------------------------------------------------
SCORE BREAKDOWN
--------------------------------------------------------------------
Quantitative Component: [XX]/55
- IT Company Experience:      [X]/15
- Big 4/International Firm:   [X]/10
- International Education:    [X]/8
- English Proficiency:        [X]/8
- IP/IT Experience:           [X]/7
- International Transactions: [X]/7
Qualitative Component: [XX]/45
- Tech Industry Interest:     [X]/8
- Business Orientation:       [X]/8
- Growth Mindset:             [X]/8
- Adaptability:               [X]/7
- Proactivity:                [X]/7
- Quantified Results:         [X]/7
```

```
--------------------------------------------------------------------
KEY STRENGTHS
--------------------------------------------------------------------
[Bullet points identifying 2-4 primary strengths with CV evidence]
--------------------------------------------------------------------
AREAS OF CONCERN
--------------------------------------------------------------------
[Bullet points identifying 1-3 concerns or gaps, if any]
--------------------------------------------------------------------
NARRATIVE ASSESSMENT
--------------------------------------------------------------------
[2-3 paragraph summary integrating quantitative and qualitative findings,
comparing to reference profiles where relevant, and providing context
for the recommendation]
--------------------------------------------------------------------
SUGGESTED INTERVIEW FOCUS AREAS
--------------------------------------------------------------------
[2-3 specific topics to explore in interview based on CV gaps or
ambiguities identified during assessment]
====================================================================
DISCLAIMER: This assessment is a decision-support tool only. Final
hiring decisions must be made by qualified human reviewers considering
the full candidate profile and organizational needs.
====================================================================
```

**— END SYSTEM PROMPT —**

***Implementation Notes***

For replicating, the following points need to be noted. The reference profiles in section 5 might need to be updated if the model is to be applied to a significantly differing population and environment. The scoring weights in section 3 are designed for the respective data and need to be adapted to differently applied environments, namely legal markets and company structures.

# APPENDIX C. SAMPLE SCREENING OUTPUTS

This appendix offers additional examples of GPT-based screening tool results to demonstrate the variability of assessments based upon the profiles of the candidates. Figure 1 in Chapter 5 demonstrated an example score of 72/100, labeled "Strongly Recommend." The following examples will demonstrate scores based upon each category to assist the practitioner in preparing their expectations regarding results of the screening tool process. All examples will be taken from the anonymized candidate profiles in the validation dataset.

## C.1 High Score Example: 85/100 (Strongly Recommend)

This output represents an exceptional candidate profile with strong alignment across both quantitative and qualitative dimensions.

```
════════════════════════════════════════════════════════════════
CANDIDATE SCREENING ASSESSMENT
════════════════════════════════════════════════════════════════

OVERALL SCORE: 85/100
RECOMMENDATION: STRONGLY RECOMMEND

────────────────────────────────────────────────────────────────
SCORE BREAKDOWN
────────────────────────────────────────────────────────────────
Quantitative Component: 48/55
  - IT Company Experience:      15/15  (6 years at tech companies)
  - Big 4/International Firm:   10/10  (Deloitte Legal, 2 years)
  - International Education:     8/8   (LLM, University of Edinburgh)
  - English Proficiency:         8/8   (C1, IELTS 8.0)
  - IP/IT Experience:            7/7   (Data protection lead)
  - International Transactions:  0/7   (Limited evidence)

Qualitative Component: 37/45
  - Tech Industry Interest:      8/8   (Tech blog author, speaker)
  - Business Orientation:        8/8   (Revenue impact quantified)
  - Growth Mindset:              8/8   (CIPP/E, CIPM, continuous dev)
  - Adaptability:                7/7   (3 sector transitions)
  - Proactivity:                 3.5/7 (Some initiative evidence)
  - Quantified Results:          2.5/7 (Limited metrics beyond certs)

────────────────────────────────────────────────────────────────
KEY STRENGTHS
────────────────────────────────────────────────────────────────
• Exceptional IT sector depth with 6 years across three technology
  companies of increasing scale and complexity
• Strong international credentials: UK LLM, Big 4 experience, and
  certified English proficiency position candidate well for global
  operations
• Demonstrated thought leadership in privacy/data protection space
  through published articles and conference presentations
• Clear pattern of continuous professional development with multiple
  IAPP certifications obtained over three-year period

────────────────────────────────────────────────────────────────
AREAS OF CONCERN
────────────────────────────────────────────────────────────────
• Limited evidence of cross-border transactional work despite
  international background; experience appears primarily advisory
• Quantified business impact described primarily in terms of
```

```
  compliance metrics rather than commercial outcomes

────────────────────────────────────────────────────────────────
NARRATIVE ASSESSMENT
────────────────────────────────────────────────────────────────
This candidate presents an exceptionally strong profile for in-house
legal roles in the technology sector. The combination of substantial
IT company experience (6 years), prestigious international credentials
(UK LLM, Big 4 background), and specialized expertise in data
protection creates a compelling qualification set.

The candidate's genuine engagement with technology—evidenced by
authored content and speaking engagements—suggests strong cultural
fit with innovative tech environments. The progression from Big 4
consulting to in-house roles demonstrates successful adaptation to
different professional contexts.

Profile characteristics align closely with Reference Profile HS-1
(Score: 87), which achieved successful long-term retention. The
primary gap relative to that profile is limited M&A/transactional
experience, which may be less critical depending on role requirements.

────────────────────────────────────────────────────────────────
SUGGESTED INTERVIEW FOCUS AREAS
────────────────────────────────────────────────────────────────
• Explore appetite for transactional work and commercial deal support
  beyond advisory/compliance functions
• Discuss experience working with product and engineering teams to
  assess cross-functional collaboration depth
• Clarify motivations for current transition and long-term career
  trajectory expectations

════════════════════════════════════════════════════════════════
DISCLAIMER: This assessment is a decision-support tool only. Final
hiring decisions must be made by qualified human reviewers.
════════════════════════════════════════════════════════════════
```

### C.2 Borderline Example: 49/100 (Consider with Caution)

This output represents a candidate with mixed indicators – some promising signals offset by notable gaps. Such profiles require careful human judgment.

```
════════════════════════════════════════════════════════════════
CANDIDATE SCREENING ASSESSMENT
════════════════════════════════════════════════════════════════

OVERALL SCORE: 49/100
RECOMMENDATION: CONSIDER WITH CAUTION

────────────────────────────────────────────────────────────────
SCORE BREAKDOWN
────────────────────────────────────────────────────────────────
Quantitative Component: 24/55
  - IT Company Experience:       5/15  (2 years at software company)
  - Big 4/International Firm:    0/10  (No Big 4/international)
  - International Education:     0/8   (Domestic education only)
  - English Proficiency:         5/8   (B2-C1, self-assessed)
  - IP/IT Experience:            7/7   (Software licensing focus)
```

```
  - International Transactions:  7/7   (Cross-border SaaS deals)

Qualitative Component: 25/45
  - Tech Industry Interest:      8/8   (Voluntary move to tech)
  - Business Orientation:        4/8   (Some commercial awareness)
  - Growth Mindset:              4/8   (Recent certification pursuit)
  - Adaptability:              3.5/7   (One major transition)
  - Proactivity:               3.5/7   (Process improvement example)
  - Quantified Results:          2/7   (Limited metrics provided)
```

---

KEY STRENGTHS

---

- Genuine interest in technology sector demonstrated through voluntary career pivot from traditional law firm practice
- Relevant subject matter expertise in software licensing and SaaS commercial agreements
- Evidence of proactive process improvement: implemented contract template system that CV indicates reduced turnaround time

---

AREAS OF CONCERN

---

- Limited overall IT sector experience (2 years) may indicate insufficient exposure to tech company operations and culture
- No international firm experience or education; English level self-assessed without certification—may require verification
- Career progression shows 4 employers in 6 years with average tenure of 18 months; pattern warrants exploration
- Quantified achievements largely absent from CV despite 6 years of experience

---

NARRATIVE ASSESSMENT

---

This candidate presents a mixed profile with both promising indicators and notable concerns. The voluntary transition from traditional law firm practice to technology suggests genuine sector interest, and the software licensing expertise is directly relevant. However, the limited IT experience duration (2 years) and absence of international credentials create uncertainty about readiness for senior in-house responsibilities.

The employment pattern (4 employers in 6 years) could reflect either positive adaptability or concerning instability—this ambiguity cannot be resolved from CV review alone. Similarly, the self-assessed English level (B2-C1) without certification introduces uncertainty for roles requiring regular international communication.

Profile characteristics show partial overlap with Reference Profile MS-3 (Score: 58, Outcome: Unsuccessful at 18 months), particularly regarding the experience level and tenure patterns. However, this candidate demonstrates stronger tech sector affinity than that reference profile. Outcome probability is genuinely uncertain.

---

```
SUGGESTED INTERVIEW FOCUS AREAS
────────────────────────────────────────────────────────────────
• Conduct English assessment or request certification evidence
• Explore reasons for employment transitions; distinguish between
  growth-driven moves and performance/fit issues
• Probe depth of IT industry understanding beyond licensing specialty
• Assess readiness for broad in-house remit vs. specialist role

════════════════════════════════════════════════════════════════
DISCLAIMER: This assessment is a decision-support tool only. Final
hiring decisions must be made by qualified human reviewers.
════════════════════════════════════════════════════════════════
```

### C.3 Low Score Example: 34/100 (Do Not Recommend)

This output represents a candidate whose profile shows limited alignment with success predictors. The tool recommends against proceeding but notes specific gaps that could be addressed through targeted development.

```
════════════════════════════════════════════════════════════════
CANDIDATE SCREENING ASSESSMENT
════════════════════════════════════════════════════════════════

OVERALL SCORE: 34/100
RECOMMENDATION: DO NOT RECOMMEND

────────────────────────────────────────────────────────────────
SCORE BREAKDOWN
────────────────────────────────────────────────────────────────
Quantitative Component: 15/55
  - IT Company Experience:       0/15  (No IT sector experience)
  - Big 4/International Firm:    0/10  (No Big 4/international)
  - International Education:     0/8   (Domestic education only)
  - English Proficiency:         2/8   (B1-B2, no certification)
  - IP/IT Experience:            6/7   (Some IP work mentioned)
  - International Transactions:  7/7   (Export control matters)

Qualitative Component: 19/45
  - Tech Industry Interest:      0/8   (No tech indicators)
  - Business Orientation:        4/8   (Some commercial language)
  - Growth Mindset:              4/8   (Bar admission, some training)
  - Adaptability:                0/7   (Single sector career)
  - Proactivity:                 4/7  (Committee participation)
  - Quantified Results:          7/7  (Strong metrics provided)

────────────────────────────────────────────────────────────────
KEY STRENGTHS
────────────────────────────────────────────────────────────────
• Strong quantification of achievements: CV includes specific
  deal values, case outcomes, and efficiency improvements
• International transaction exposure through export control and
  cross-border regulatory compliance work
• Professional engagement through bar association committee work
  suggests commitment to legal profession

────────────────────────────────────────────────────────────────
AREAS OF CONCERN
────────────────────────────────────────────────────────────────
• Zero IT/technology sector experience across 9-year career;
```

```
  exclusively government and heavy industry background
• No indicators of technology interest, awareness, or affinity
  anywhere in CV content
• Limited English proficiency (B1-B2) would constrain effectiveness
  in international technology company context
• Career entirely within single sector (energy/industrial)
  raises questions about adaptability to fast-paced tech culture
• CV language predominantly procedural/regulatory; limited evidence
  of business partnership orientation

────────────────────────────────────────────────────────────────
NARRATIVE ASSESSMENT
────────────────────────────────────────────────────────────────
This candidate presents a profile with limited alignment to the
predictors of success in IT company legal roles. Despite 9 years of
legal experience and some notable achievements in the energy sector,
the complete absence of technology industry exposure is a significant
gap. The model's training data indicates that candidates with zero IT
experience have substantially lower success probability regardless of
total experience level.

The candidate's CV shows no evidence of technology interest—no tech
clients, no relevant certifications, no indication of awareness of
technology sector legal issues. Combined with the limited English
proficiency and exclusively regulatory/compliance career focus, this
suggests potential cultural misalignment with fast-paced technology
environments.

Profile characteristics align closely with Reference Profile LS-2
(Score: 38, Outcome: Rejected at screening), which similarly showed
strong credentials in a non-tech sector without technology affinity
indicators. Historical data suggests low probability of success.

────────────────────────────────────────────────────────────────
SUGGESTED INTERVIEW FOCUS AREAS
────────────────────────────────────────────────────────────────
• If proceeding despite recommendation, explore motivations for
  seeking technology sector transition at this career stage
• Assess genuine interest in technology beyond job seeking
• Evaluate adaptability potential through discussion of how
  candidate has handled change in current career context
• Verify English proficiency through conversational assessment

NOTE: This candidate may be better suited for legal roles in
energy, industrial, or regulatory sectors where existing
expertise would be directly applicable.

════════════════════════════════════════════════════════════════
DISCLAIMER: This assessment is a decision-support tool only. Final
hiring decisions must be made by qualified human reviewers.
════════════════════════════════════════════════════════════════
```

### C.4 Interpretation Guidance

The above sample outputs illustrate several important aspects of score interpretation that the practitioner should recognize in interpreting the results.

***Score Distribution Expectations***

Based on the results of the validation findings described in Chapter 5, the practitioner should expect the following distribution in screening the candidate pools similar to the training data: a percentage of candidates scoring 70 and above (Strongly Recommend) ranging from 15% to 20%, a percentage scoring 55 to 69 (Recommend) ranging from 25% to 30%, a percentage scoring 40 to 54 (Consider with Caution) ranging from 25% to 30%, and a percentage scoring lower than 40 (Do Not Recommend) ranging from 25% to 30%. Major deviations in the distribution of the target population may be an indication of an irregular candidate pool or possible model drift, which indicates the need of recalibration.

***Component Score Patterns***

Comparing the examples illustrated above, it becomes clear that those scoring well usually show strengths in terms of both the quantitative and qualitative aspects, while those who score low demonstrate weaknesses in the two categories too. However, the most difficult scenarios to interpret regarding the values of the variables will present asymmetric distributions of their qualities, and this will be the borderline scenario (C.2), where the best qualitative indicators will be combined with the weak quantitative aspects, and vice versa. These cases will need the most careful human assessment because the model's linear combination will tend to neglect important interaction terms.

***Narrative Assessment Value***

The narrative assessment section provides non-numerical information that cannot be captured in a score. Practitioners should focus their attention to the comparisons to the reference profiles, where much of the assessment is grounded in historical outcomes. Note the point of uncertainty and ambiguity, where the borderline assessment is said to be uncertain in terms of employment patterns. This is where human assessments are most crucial.

***Appropriate Use of "Do Not Recommend"***

The output "Do Not Recommend," according to C.3, should not be taken to mean a definitive "no." The system points out those applicants whose profiles demonstrate a statistical tendency to correlate with a lower success rate; but discrepancies will occur in the results for the individual applicants in relation to the trends that the data tends to follow. The system is impossible to validate in terms of what might be omitted regarding a CV and the data it contains, such as an applicant's performance in an interview, his references, a particular job description, and the organization's demands. Companies with special obligations, e.g., regulatory expertise in a highly compliant function, will be able to nominate those people for the job to whom the model has assigned a low score to its indicators of success in the first place. The proposed areas of focus for the interviews in these instances are relevant in identifying the type of information that might justify proceeding given the low score.

***Limitations Illustrated***

The examples also illustrate the natural constraints of the tool in the sense that the model has no capability to assess interpersonal competencies, cultural fit beyond proxy indicators, and potential for development. For instance, a candidate such as in example C.3, who has excellent quantification skills and engagement with the profession, has a high chance of success if there is a clear motivation to switch sectors, but this motivation cannot be determined from the data in the CV examples. Human judgment remains essential for all hiring decisions.

## REFERENCES


Becker, G. S. (1964). Human capital: A theoretical and empirical analysis, with special reference to education. National Bureau of Economic Research.

Becker, G. S. (1975). Human capital: A theoretical and empirical analysis, with special reference to education (2nd ed.). National Bureau of Economic Research.

Becker, G. S. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education* (3rd ed.). University of Chicago Press.

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science, 358*(6370), 1530–1534. **https://doi.org/10.1126/science.aap8062**

Cappelli, P., & Keller, J. R. (2014). Talent management: Conceptual approaches and practical challenges. *Annual Review of Organizational Psychology and Organizational Behavior, 1*(1), 305–331. **https://doi.org/10.1146/annurev-orgpsych-031413-091314**

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. **https://doi.org/10.1177/001316446002000104**

European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). *Official Journal of the European Union, L*119, 1–88. **https://eur-lex.europa.eu/eli/reg/2016/679/oj**

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29–36. **https://doi.org/10.1148/radiology.143.1.7063747**

Hoffman, M., Kahn, L. B., & Li, D. (2018). Discretion in hiring. *The Quarterly Journal of Economics, 133*(2), 765–800. **https://doi.org/10.1093/qje/qjx042**

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). John Wiley & Sons. **https://doi.org/10.1002/0471722146**

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Sage Publications.

JRG Partners. (2024). Data-driven recruitment strategies to hire top talent. JRG Partners. **https://www.jrgpartners.com/data-driven-recruitment-strategies-to-hire/**

Kristof, A. L. (1996). Person-organization fit: An integrative review of its conceptualizations, measurement, and implications. *Personnel Psychology, 49*(1), 1–49. **https://doi.org/10.1111/j.1744-6570.1996.tb01790.x**

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. **https://doi.org/10.1111/j.1744-6570.2005.00672.x**

Kuncel, N. R., Klieger, D. M., Connelly, B. S., & Ones, D. S. (2013). Mechanical versus clinical data combination in selection and admissions decisions: A meta-analysis. *Journal of Applied Psychology, 98*(6), 1060–1072. **https://doi.org/10.1037/a0034156**

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. **https://doi.org/10.2307/2529310**

Lievens, F., & Sackett, P. R. (2012). The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology, 97*(2), 460–468. **https://doi.org/10.1037/a0025741**

Lviv IT Cluster. (2023). *IT Research Ukraine 2023*. **https://itcluster.lviv.ua/wp-content/uploads/2023/12/it-research-ukraine-2023-public-eng.pdf**

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology, 79*(4), 599–616. **https://doi.org/10.1037/0021-9010.79.4.599**

OpenAI. (2025). Introducing GPT-5.2. **https://openai.com/index/introducing-gpt-5-2/**

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. **https://doi.org/10.1037/apl0000994**

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262–274. **https://doi.org/10.1037/0033-2909.124.2.262**

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics, 87*(3), 355–374. **https://doi.org/10.2307/1882010**

Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review, 61*(4), 15–42. **https://doi.org/10.1177/0008125619867910**

Ton, Z., & Huckman, R. S. (2008). Managing the impact of employee turnover on performance: The role of process conformance. *Organization Science, 19*(1), 56–68. **https://doi.org/10.1287/orsc.1070.0294**

Van Iddekinge, C. H., & Ployhart, R. E. (2008). Developments in the criterion-related validation of selection procedures: A critical review and recommendations for practice. *Personnel Psychology, 61*(4), 871–925. **https://doi.org/10.1111/j.1744-6570.2008.00133.x**