

RESEARCH PAPER | ENGLISH EDITION | OCTOBER 2026

WHO AUDITS THE GATEKEEPERS?

Hiring Governance and the Quality of Early Rejection Decisions

Research essay and practical framework for governing the quality of early-stage screening

Serikzhan Umarov

Independent Research | U Cool Research

Original Russian edition: September 2026

CORE IDEA

Recruitment should not be a wall between professionals. It should be infrastructure that helps them find one another. The authority to make an early rejection decision should have a verifiable basis.

Executive Framing

THE PROBLEM

A hiring manager sees the candidates who were allowed to advance, but usually has little visibility into the quality of the rejected-candidate pool. A weak candidate who is mistakenly advanced can often be identified later in the process. A strong candidate who is mistakenly screened out at the beginning may remain invisible.

THE HYPOTHESIS

Some early professional rejections may not be reproducible when independently reviewed by a domain reviewer.

THE PROPOSED RESPONSE

The Hiring Governance Framework separates objective eligibility checks, judgments of professional relevance and the final decision. Ambiguous professional rejections are escalated for domain review and leave an auditable decision trail.

THE TEST

A 30-Day Diagnostic Pilot: Baseline → Blind Re-review → Root-Cause Analysis → Controlled Change → Follow-up Sample → Decision.

The Candidate the Hiring Manager Never Sees

On Monday morning, the head of a finance function opens the calendar. Three interviews are scheduled for the week. The recruiter has sent over the “best candidates.”

The manager reviews their CVs, asks questions, compares experience and, by the end of the week, concludes: “Not a bad pipeline.”

One person is missing from this picture.

That person applied too.

Perhaps the candidate had relevant experience, an unusual but strong combination of Finance + Analytics + Technology, a thoughtful question in the cover letter and the ability to solve the very problem for which the role had been opened.

But the application ended earlier - during initial screening.

The hiring manager will never see the CV. The team will never ask the candidate a single question. The organization may never discover that it could have made a mistake.

A weak candidate who is mistakenly allowed to advance will usually be detected later. A strong candidate who is mistakenly stopped at the beginning may never be detected at all.

This is the invisible false negative.

That is why the central question of this research is not whether HR professionals are “good” or whether recruiters are necessary.

The question is more difficult: Who audits the quality of the authority to make an early rejection decision?

Who Audits the Gatekeepers?

How opaque initial screening can turn recruitment into an uncontrolled filter - and how to build an auditable hiring system

Long-form research essay | English Edition | October 2026

Abstract

Modern hiring contains an underexamined asymmetry of decision authority.

A functional leader may be looking for a strong finance professional, accountant, engineer, physician, analyst or software developer, yet before two professionals ever speak to one another, a third participant often stands between them: the recruiter.

The recruiter may be the first person given the authority to stop an application. At the same time, that person's depth of domain expertise may be lower than that of the hiring manager, while the basis for rejection may remain invisible to the candidate, the manager and any independent control function.

This paper does not argue that HR or recruitment is unnecessary as a profession. Professional recruitment performs essential functions: sourcing, coordination, checking objective constraints, organizing interviews, supporting candidate communication, documentation and assisting the hiring manager.

The critique concerns a different decision architecture: one in which an intermediary function quietly becomes an opaque authority over professional access.

The central proposition is this: If an organization allows one participant to reject a candidate before that candidate has contact with an appropriate professional reviewer, the organization should be able to demonstrate that the decision was based on pre-defined, job-related criteria rather than personal impression, superficial reading, convenience, stereotype or informal preference.

Otherwise, an invisible false negative may occur: a potentially suitable candidate disappears from the hiring funnel, while the organization may never learn that the decision was questionable.

The paper brings together three layers of material: more than seventeen years of personal observation of digital job-search and application processes through hh.kz; research evidence on discrimination, bias and the validity of personnel-selection methods; and a proposed Hiring Governance Framework that can be tested empirically within an organization.

The framework introduces a Recruitment Audit Trail, separates decision rights between recruiters and appropriate professional reviewers, uses the Four-Eyes Principle and Random Rejection Audit, applies structured reason codes, and proposes diagnostic measures including False Negative Review Rate (FNRR) and Domain Override Rate (DOR).

The objective is not to eliminate recruitment or to substitute one person's judgment for another's. It is to make consequential early screening decisions sufficiently structured, reconstructable and auditable that an organization can test their quality rather than simply assume it.

Keywords: *Hiring Governance; Decision Quality; Recruitment; Initial Screening; Invisible False Negatives; Auditability; Personnel Selection; Decision Rights; Kazakhstan*

1. Not a War on HR, but a Governance Question

Emotional criticism of recruitment does not arise from nowhere.

For a candidate, a rejection is not an abstract status in an applicant tracking system. It represents a lost opportunity, time and preparation, sometimes months of searching, and potentially a real effect on income and career trajectory. That is why a generic response after a demanding application process can feel particularly frustrating.

But if the objective is to improve the system, it is not enough to say that a particular recruiter was inattentive or incompetent. The question has to be framed in a way that can be addressed through governance, data and control.

The central question is: Why does an organization delegate to one intermediary the authority to terminate a potential professional connection between a candidate and the future team when the quality of that decision is often not measured?

In financial control, “I looked at it carefully” would not normally be considered sufficient evidence. A payment may require defined authority, supporting documentation, segregation of duties, an audit trail and approval. In data engineering, lineage matters. In software engineering, code review and change history matter.

Yet in hiring, a decision that may deprive an organization of a valuable professional can sometimes be reduced to a single click - “Reject” - followed by a polite template message.

This is not an argument for adding bureaucracy to every application. It is an argument for making control proportionate to the cost of error.

The more professionally complex the role, the greater the risk of allowing someone without sufficient domain expertise to exercise final veto power before the candidate has been seen by someone capable of evaluating that professional background.

The object of this research is therefore not “bad HR” as a category. It is the decision architecture.

That architecture needs to distinguish between the legitimate and valuable role of the recruiter and an excessive concentration of decision rights.

A well-designed system should protect the candidate from arbitrary treatment, the hiring manager from the invisible loss of strong candidates, the organization from discrimination and reputational risk, and professional recruiters themselves from accusations that cannot be confirmed or disproved because the underlying decision data do not exist.

2. More Than Seventeen Years of Observation: What a Candidate Can See on hh.kz

My starting point is not a single rejection.

It is an observation accumulated over more than seventeen years of posting and updating CVs, submitting applications, receiving invitations and rejections, and obtaining employment through hh.kz.

That experience should not be presented as representative statistics for the Kazakhstan labor market. It is one long-running case.

Its duration, however, makes it useful as a source of hypotheses.

Certain recurring patterns will be familiar to many job seekers. An application may remain unopened for weeks and the vacancy then closes. A cover letter may contain a specific question, yet the eventual response is generic and leaves the question unanswered. A CV may be opened and followed quickly by a rejection stating that the employer has “carefully reviewed” the application. In other cases, the candidate may see repeated views but receive no substantive explanation.

There are also cases in which the communication is clearly professional: the recruiter answers questions, asks for clarification where the profile is ambiguous, follows up with feedback, and recognizes when an issue requires the involvement of the hiring manager.

It is this contrast that creates the research question.

The hh.kz documentation confirms an important part of the process mechanics. After a candidate applies, the employer receives summary information from the CV and the cover letter, if one has been provided.

The employer can configure evaluation parameters and, when reviewing the application, choose actions such as “Invite” or “Reject.” Status information may allow the candidate to see whether the CV has been viewed, but it does not reveal the quality of that review, the actual amount of reading time, the criteria applied, or who participated in the decision [1].

This creates an important limitation.

It would not be valid to say, for example: “The CV was open for two minutes, therefore it was not read.”

A platform interface is not a stopwatch for cognitive effort. A late-night timestamp likewise does not establish where the reviewer was located or how seriously the application was considered.

But the underlying governance question remains.

If an organization tells a candidate that the application was “carefully reviewed,” can the organization itself reconstruct which job-related criteria were actually assessed, who assessed them, and why the conclusion was negative?

If it cannot, the word “carefully” is not evidence of the process. It is part of a communication template.

For that reason, my experience is used here not as a verdict on the industry, but as a generator of testable hypotheses:

- Is there a meaningful proportion of applications that are processed only after a substantial delay?
- Are cover letters and candidate questions actually read?
- How often does an early professional rejection occur without domain review?
- How often would a hiring manager reverse an initial rejection in a blind re-review?
- Does the quality of initial screening differ materially between recruiters?
- Is the quality of candidate communication associated with the quality of the resulting shortlist?

3. The Gatekeeper: When an Intermediary Controls Access

In a conventional hiring funnel, the recruiter sits between the candidate market and the hiring manager.

This arrangement is reasonable. A hiring manager cannot personally process hundreds of applications.

The governance problem begins when saving the manager’s time means transferring not only administrative work to the recruiter, but also the authority to make a final interpretation of professional suitability.

Consider a Senior Financial Analyst position.

A CV may contain IFRS, FP&A, budgeting, cash flow, Power Query, DAX, SQL, financial modelling, audit, data analytics and experience with automation.

For an experienced CFO or Head of FP&A, these elements may form a coherent professional profile. The manager may be able to see where the capabilities reinforce one another, where adjacent experience is transferable, where a gap could realistically be closed, and where an impressive keyword does not represent sufficient depth.

For a generalist recruiter, the same elements may instead become a checklist.

This creates an epistemic asymmetry: a person with less domain expertise may be given the authority to decide whether someone with greater expertise in that domain should have a conversation with the person capable of assessing it.

A strong recruiter responds to this problem by learning the domain, calibrating with the hiring manager, defining criteria in advance, distinguishing must-have from trainable requirements, and recognizing uncertainty when it exists.

That is precisely why professionalism should not be expressed as unlimited discretion. It should be expressed through the ability to work within a reproducible decision process.

The key principle is: “I cannot confidently establish that this candidate is relevant” does not mean “this candidate is not relevant.”

For a complex role, uncertainty should trigger escalation to a domain reviewer rather than automatically produce rejection.

4. What the Research Shows: Discrimination Can Occur Before the Interview

One of the most uncomfortable findings in the evidence on hiring comes from correspondence experiments.

In these studies, researchers create closely matched candidate profiles and randomly vary a characteristic that should not determine professional suitability. They then observe whether employers respond differently.

A classic study by Marianne Bertrand and Sendhil Mullainathan sent fictitious CVs in response to real job vacancies in Boston and Chicago. The CVs were randomly assigned names that were perceived as associated with White or African American applicants.

The authors found that CVs with names perceived as “White” received approximately 50% more employer callbacks. They also found that improvements in resume quality did not generate equal returns across the two groups [2].

This US study cannot be turned into a statistic about Kazakhstan.

What it does demonstrate, within the conditions of the experiment, is a causal mechanism: when professional information is held nearly constant, changing a socially perceived characteristic can change the probability of a positive employer response.

A meta-analysis by Quillian and colleagues combined 28 US field experiments, 55,842 applications and 26,326 job vacancies. The authors reported that White applicants received more callbacks on average than African American and Latino applicants. For African American applicants, the analysis did not find convincing evidence that the gap had materially declined over the period studied [3].

A 2023 meta-analysis published in the European Economic Review synthesized a quasi-exhaustive body of correspondence experiments conducted between 2005 and 2020 across ten potential grounds of discrimination. It found significant unequal treatment not only in relation to racial and ethnic characteristics, but also for older applicants, people with disabilities and some other groups [4].

These findings do not mean that every rejection is discriminatory.

They also do not justify attributing discriminatory intent to a particular recruiter without evidence.

Their implication is narrower but still important: the assumption that human initial screening is objectively neutral by default is not supported by the empirical evidence.

A hiring system therefore needs to be designed with the possibility of bias in mind.

Bias does not have to be conscious.

A reviewer does not need to think, “I do not want someone of this nationality” or “I do not want someone of this age.” It may be enough for the decision to rely on an unstructured impression: “Something does not feel right.” “Not really one of us.” “I do not feel the fit.”

The fewer the pre-defined criteria and independent checks, the easier it becomes for personal impression to present itself as professional judgment.

5. Structured and Unstructured Selection

The science of personnel selection has examined for decades how well different methods predict future job performance.

The classic work of Schmidt and Hunter synthesized decades of research and demonstrated the practical value of valid, structured selection methods. Among the stronger combinations were structured interviews and work-sample approaches used alongside other assessments [5].

A major review by Levashina, Hartwell, Morgeson and Campion shows that interview quality depends heavily on structure [6].

Relevant design questions include:

- Are candidates asked comparable questions?
- Are the criteria defined before the interview?
- Are responses assessed against a rubric?
- Are ratings made independently before discussion?
- Are interviewers trained?
- Are the questions related to the actual work of the role?

CIPD likewise recommends structured interviews: pre-defined questions, a consistent order, common scoring criteria and independent assessment before group discussion. It also points to the value of panel approaches and multiple sources of evidence [7].

This leads to an uncomfortable question: If organizations are willing to structure the later interview stage, why is the earlier screening stage - which can permanently remove a candidate from the process - sometimes left unstructured?

A CV cannot be assessed entirely mechanically.

Career paths are non-linear. Job titles vary across organizations. Adjacent experience may be more valuable than a formal title match. A candidate may have the required capability without using the exact language expected by a screening checklist.

But complexity is not an argument for uncontrolled intuition.

It is an argument for a transparent assessment rubric and for escalation to a domain reviewer where professional relevance cannot be established reliably at the initial screening stage.

6. Kazakhstan: Equal Opportunity Is More Than Good Practice

The Labour Code of the Republic of Kazakhstan establishes principles of freedom of labour, equality of rights and opportunities, and protection against discrimination.

Article 6 prohibits discrimination in the exercise of labour rights on grounds including origin, social status, sex, race, nationality, language, attitude toward religion, beliefs, place of residence, age, physical disability and other circumstances. People who believe they have experienced discrimination may seek protection through a court or other competent bodies.

Article 25 establishes guarantees of equal rights and opportunities when entering into an employment contract [8].

In September 2026, a Kazakhstan government resource explaining the inadmissibility of discrimination in hiring expressed a practical principle: employers should evaluate applicants on the basis of qualifications, education, professional knowledge and experience, while a refusal should be grounded in the requirements of the vacancy and the candidate's professional qualities [9].

For governance purposes, the paper expresses this logic as:

GOVERNANCE LOGIC

Role Requirement → Candidate Evidence → Assessment → Decision

This does not mean that Kazakhstan law requires an employer to provide every applicant with an internal scorecard.

The point is internal auditability.

If a complaint, audit or internal investigation arises, can the organization reconstruct why the candidate was rejected and which job-related criteria were actually applied?

The more decisions remain recorded only as “not suitable,” the more difficult it becomes to distinguish a legitimate professional judgment from arbitrary preference.

7. The Invisible False Negative - the Most Dangerous Error

Recruitment analytics often focuses on time-to-hire, cost-per-hire, source-of-hire, offer acceptance and employee retention.

But these measures mostly describe the people whom the system has seen and allowed to advance.

Consider two candidates.

A weak candidate is advanced by the recruiter to a professional interview. The hiring manager is likely to notice the mistake. The false positive becomes visible.

A strong candidate is rejected during initial CV screening. The hiring manager never sees the CV. Unless someone audits the decision, the organization may never learn that the rejection was questionable.

This is the invisible false negative.

A hiring manager may say, “The recruiter sends me good candidates,” because the manager sees only the shortlist.

But the manager does not know the quality of the rejected-candidate pool.

A system can therefore show acceptable precision among candidates who advance while still having poor recall among strong candidates in the broader applicant pool.

For an analyst, the analogy is straightforward: a classifier cannot be evaluated by observing only its positive predictions.

The same logic applies in control environments. Reviewing only successful transactions while ignoring exceptions does not tell you whether the control captures what it is supposed to capture.

This leads to a proposed diagnostic measure:

False Negative Review Rate (FNRR)

False Negative Review Rate is the proportion of audited rejections that an independent domain reviewer, during a blind re-review, would move at least into the category “Short Conversation.”

FNRR is not perfect ground truth.

A domain reviewer can also be wrong. Professional judgment can differ between reviewers. But regular sampled re-review makes an otherwise invisible disagreement observable and measurable.

8. “We Carefully Reviewed Your Application”: Language Without Evidence

A standardized rejection message is not inherently a problem.

When hundreds of people apply for a role, it is unrealistic to expect an individually written essay for every candidate.

The problem begins when a template makes a claim about the quality of the underlying process that the organization itself cannot verify.

What does “carefully reviewed” actually mean?

- Was the cover letter read?
- Were the candidate's questions noticed?
- Were achievements compared with the must-have criteria?
- Was adjacent or transferable experience considered?
- Did a domain reviewer see the materials?
- Was the rejection based on an objective constraint, or on a subjective impression?

If the answers are unknown, polite wording does not constitute evidence of a controlled process.

More honest communication may be shorter and more precise.

But the primary object of reform is not the wording of the rejection message.

It is the data and decision process behind it.

9. Recruiter vs Hiring Manager: Separating Decision Rights

Eliminating recruitment and requiring hiring managers to read every application would be a poor universal strategy.

A more practical objective is to separate different types of decisions.

A recruiter should have strong decision authority where the criterion is objective and pre-defined, for example:

- a legally required licence or certification;
- location or working arrangement;
- availability or potential start date;
- compensation range;
- a required language, where genuinely necessary for the role;
- clear absence of an unambiguous must-have requirement.

By contrast, the interpretation of complex professional relevance should belong to the hiring manager or an appropriate domain reviewer.

Questions in this category include:

- How transferable is the candidate's experience?
- Are different technologies functionally equivalent?
- Could an unusual career path be an advantage rather than a weakness?
- Can a gap be compensated for by adjacent capabilities?

The principle is: A professional rejection should rest on verifiable professional criteria.

For senior and specialized roles, an ambiguous professional rejection should not become final without domain review.

10. Recruitment Audit Trail

If hiring is a business process, it should have traceability.

A Recruitment Audit Trail should make it possible to reconstruct the path from application to decision:

DECISION TRACE

Application ID → Vacancy Version → CV Version → Cover Letter → Candidate Questions → Screening-Criteria Version → Recruiter Assessment → Domain Review → Decision → Reason Code → Communication → Timestamp

Versioning matters.

If the vacancy requirements or screening criteria change during the hiring process, the organization should be able to identify which version governed a particular decision.

Otherwise, old decisions can later be explained using criteria that did not exist when those decisions were actually made.

“Not suitable” is an almost useless reason code.

A more auditable structure should distinguish between reasons such as:

- legally required qualification absent;
- mandatory certification or licence absent;
- a specific must-have experience criterion not demonstrated;
- compensation mismatch;
- location or availability mismatch;
- duplicate application or candidate withdrawal;
- vacancy closed or placed on hold;
- domain review required.

A particularly important category is: Uncertain / Domain Review

This category institutionalizes professional humility.

A recruiter should not be penalized for recognizing that a complex profile cannot be assessed confidently. In a controlled process, escalation is the correct action when the evidence is insufficient for a reliable professional judgment.

The audit trail should also preserve candidate questions and whether those questions were answered.

Access should be role-based.

Auditability does not mean making candidate information public. It means that authorized participants can reconstruct and inspect the decision process when necessary.

11. Four-Eyes Principle and Random Rejection Audit

For selected categories of roles, a negative professional decision should receive a second qualified review.

This does not mean that the CFO must read every CV.

Organizations can designate domain reviewers: for example, a senior accountant, lead analyst, head of an engineering function or another appropriately qualified professional.

Four-Eyes Principle

The Four-Eyes Principle adds a second qualified reviewer where a professional rejection is sufficiently complex, ambiguous or consequential to justify another perspective.

Random Rejection Audit

At regular intervals, a random sample of early rejections can be assigned to independent domain reviewers.

Where practical, the original decision and the identity of the recruiter should be concealed to reduce anchoring.

The reviewer may classify each case into categories such as:

- Reject
- Short Conversation
- Domain Interview
- Strong Candidate

Suppose that, out of 100 audited rejections, domain reviewers consider 20 worthy of at least a conversation.

That does not automatically mean that the organization has discovered 20 proven errors.

It does, however, create a strong signal for root-cause analysis.

The same principle can be applied to the shortlist. Auditing both rejected and advanced candidates can provide a more balanced picture of possible false negatives and false positives.

A partially blinded review may also be tested where lawful and operationally feasible - for example, by concealing a photograph, age, name or another characteristic unrelated to the job.

This is not a universal solution.

It is an experimental control that can help test whether outcomes change when selected identifiers are removed.

12. Metrics That Make the Disagreement Testable

The purpose of the following measures is not to create a league table of “good” and “bad” recruiters.

Their purpose is to turn otherwise invisible disagreements into observable process evidence.

False Negative Review Rate (FNRR)

The proportion of audited rejections that a domain reviewer, on re-review, moves at least into the “Short Conversation” category.

Domain Override Rate (DOR)

The proportion of cases in which a domain reviewer changes the recruiter’s original professional assessment.

Reason-Code Completeness (RCC)

The proportion of decisions with a specific recorded rationale linked to a vacancy criterion.

Time-to-First-Review (TFR)

The time between application submission and the first documented review.

Time-to-Decision

The time between application submission or first review, depending on the organization’s definition, and the recorded decision. This is an operational-discipline measure rather than an indicator of decision quality by itself.

Unanswered Candidate Question Rate

The proportion of substantive candidate questions that remain unanswered where the process is intended to provide a response.

Rejection Concentration by Reviewer

Differences in rejection rates between reviewers after accounting for the vacancy mix. This is a calibration signal, not a ranking of “bad recruiters.”

Candidate Reconsideration Rate

The proportion of applications returned to the hiring process after re-review.

Adverse Outcome Monitoring

Monitoring differences in hiring-process outcomes between groups where collecting and analyzing such data is lawful, ethical and appropriately governed. This requires privacy governance, control of personal data and appropriate legal review.

These metrics should be used to diagnose and improve the process, not as simplistic punishment mechanisms.

The central question is: Which errors in initial screening can the organization detect, reproduce and reduce?

13. An Experiment Almost Any Company Can Run

Select 5-10 completed professional vacancies.

Randomly sample 100-300 CVs that were rejected before a domain interview.

Remove the original decision, the recruiter's identity and, where lawful and practical, identifiers that are not relevant to the job.

Give the vacancy criteria and the CVs to two independent domain reviewers.

Each reviewer assesses the applications independently using categories such as:

- Reject
- Short Conversation
- Interview

The reviewer should also record a confidence level and a reason code.

The results can then be compared with the original initial screening decisions.

Where the decisions differ, the organization can examine possible causes, including:

- insufficient domain knowledge;
- an overly rigid keyword filter;
- poorly defined vacancy requirements;
- workload or screening overload;
- inconsistent criteria;
- risk of bias;
- incomplete information.

The process can then be repeated after a targeted change.

If the False Negative Review Rate declines without a disproportionate increase in managerial workload, that provides evidence that the change is improving the process.

If FNRR is low from the outset, the recruitment function gains empirical support for the quality of its early screening decisions.

In either case, the organization learns something it could not learn by looking only at the shortlist.

This is what distinguishes a governance approach from public blame.

It does not assume guilt.

It requires auditability.

14. Algorithms and AI Do Not Automatically Solve the Problem

There is an understandable temptation to address human subjectivity with applicant-tracking filters and AI.

But automation does not eliminate the governance question. It relocates it.

- Who defined the features?
- What data was the model trained on?
- What counts as a “successful employee”?
- How is adverse impact measured?
- How can a human review or challenge the result?

US guidance on the use of algorithms in hiring warns that hiring technologies can unfairly screen out qualified people with disabilities and that employers should evaluate such tools before and during their use [10].

That guidance relates to the United States, not Kazakhstan.

But the broader engineering implication remains relevant: An automated screening stage also requires controls.

AI can support the process. It can help to:

- extract skills from application materials;
- match evidence against vacancy criteria;
- detect candidate questions that may have been overlooked;
- suggest reason codes;
- identify cases for audit.

But AI should not become an even less transparent gatekeeper.

GOVERNING PRINCIPLE

AI proposes. An accountable person makes the decision. Evidence is retained. Exceptions are subject to review.

15. Candidate Experience as an Indicator of Process Quality

Courtesy is not the same thing as selection quality.

But poor communication can be a signal of weakness in the underlying operating process.

CIPD notes that a poor interview experience can damage the employer brand, while timely follow-up and feedback can improve the candidate experience [7].

If candidate questions are repeatedly lost, statuses remain unchanged for months, vacancies close without a clear conclusion, or promised feedback does not arrive, the explanation may lie in:

- excessive workload;
- weak process ownership;
- unclear service levels;
- poor data discipline.

None of these observations proves that a screening decision was wrong.

They do, however, provide operational signals that the process may not be well controlled.

At very high application volumes, individualized analysis of every CV is unrealistic.

An organization can still define clear service levels:

- when an application is considered reviewed;
- which candidate questions require a response;
- at which stages individualized feedback is provided;

- when a candidate should receive a status update;
- how vacancies placed on hold are handled.

The problem is not the existence of a standardized message.

The problem is when standardized communication conceals the absence of a managed process.

16. What Should Not Be Claimed Without Evidence

Strong criticism requires self-restraint.

A timestamp cannot establish that a recruiter reviewed a CV “from a café, a shisha lounge or while lying at home.”

A recruiter’s age or sex cannot be used to infer professional incompetence.

Without evidence, it is not valid to claim that “most recruiters do not understand the profession they are hiring for.”

Statements like these weaken the evidentiary architecture of the argument.

A stronger approach is to document what is actually observable:

- the application was not reviewed before the vacancy closed;
- a candidate question remained unanswered;
- the rejection contains no job-related decision basis;
- no domain reviewer participated;
- the criteria were not documented;
- no audit trail exists;
- a blind re-review produced a different professional assessment.

If the objective is public attention, there is always a temptation to maximize outrage.

If the objective is to improve practice, the stronger objective is to maximize verifiability.

17. The Hiring Governance Framework

The proposed Hiring Governance Framework consists of seven layers.

01 - Vacancy Contract

Before the vacancy is published, the hiring manager and recruiter define the role’s must-have, trainable and nice-to-have criteria, together with the grounds for administrative rejection. The requirements should be linked to the actual work of the role.

02 - Structured Intake

Each application receives an Application ID, a vacancy version and a linked package of application materials. A cover letter and candidate questions should not become easily lost attachments or isolated messages disconnected from the application record.

03 - Eligibility Gate

The recruiter checks objective eligibility constraints and independently decides only those cases that were defined in advance.

04 - Professional Relevance Gate

Complex professional relevance is assessed against a structured rubric. Ambiguous cases are escalated to a domain reviewer. For senior and specialized roles, the Four-Eyes Principle is applied where appropriate.

05 - Structured Assessment

Interviews, work samples and case exercises are designed around job-related constructs. Assessments are made independently before the reviewers discuss their conclusions.

06 - Decision & Communication

The decision receives a reason code, an evidence reference and a decision owner. The candidate receives communication that is appropriate to the stage and circumstances of the process.

07 - Audit & Learning

Random Rejection Audit, calibration sessions, drift monitoring and candidate feedback are used to improve the process over time.

This architecture does not eliminate HR or recruitment.

It changes recruitment from a largely personal filter into a more controlled professional function whose consequential decisions can be examined and improved.

18. Who Should Be Able to See the Decisions?

Complete transparency to every employee would be a mistake.

CVs contain personal data and other sensitive information.

Access should therefore be role-based.

The hiring manager should be able to see the application materials for the manager's own vacancy and the basis for key decisions.

Domain reviewers should receive the materials necessary for the professional review they are asked to perform.

HR leadership should have access to process analytics.

Compliance and internal audit should receive controlled access where appropriate.

A governance system should not become a paperwork factory.

Reason codes should take seconds to record. A structured assessment rubric should take minutes, not hours. Re-review should be applied on a risk-based basis.

Priority cases may include:

- senior and specialized roles;
- scarce capabilities;
- ambiguous professional profiles;
- high-impact positions;
- randomly sampled audit cases.

The objective is not maximum documentation.

The objective is to make the process observable enough that weak practices cannot continue for years without being detected.

19. What the Business Can Gain

A more auditable hiring process can create several forms of potential business value.

1. Recovery of Strong Candidates

A structured review path can reduce the risk that a strong candidate is lost because of a superficial or overly narrow screen.

2. Better Vacancy Design

If reviewers repeatedly disagree about the same role, the problem may lie not with the candidates but with the vacancy criteria themselves. Repeated disagreement can reveal where requirements are unclear, internally inconsistent or poorly connected to the actual work.

3. Lower Key-Person Dependence

Hiring quality becomes less dependent on the judgment or preferences of a single recruiter or reviewer.

4. Stronger Employer-Brand Protection

A candidate may still receive a rejection, but the organization can operate a process that is more professional, consistent and defensible.

5. Better Evidence When Concerns Arise

An audit trail gives the organization a basis for reconstructing and investigating consequential decisions.

6. Development of Recruiter Expertise

Measures such as Domain Override Rate, together with calibration, can help identify where additional domain learning or reviewer alignment may be useful.

7. More Mature Recruitment Analytics

Recruitment quality is no longer assessed only by how quickly a vacancy is filled. The organization begins to ask whether the negative decisions made earlier in the funnel are themselves sufficiently reliable and auditable.

20. Who Audits the Gatekeepers?

We are accustomed to thinking of the recruiter as the person who connects a candidate with an employer.

In many organizations, however, the role is more consequential than that.

The recruiter does not merely connect the two sides.

The recruiter may also decide whether that connection happens at all.

A CFO may be looking for a strong financial analyst.

A Chief Accountant may need an IFRS specialist.

A Head of Data may be looking for a data analyst.

But before two professionals speak to one another, a third person often stands between them: a recruitment professional.

And that person may be the first to say no.

Not the CFO.

Not the Chief Accountant.

Not the Head of Analytics or Data.

Not the candidate's future colleagues.

The recruiter.

This creates a fundamental governance question: On what basis should a person who is not required to possess the same professional qualifications as the people being screened have sole authority to decide whether one professional is allowed to speak with another?

The problem is not the existence of recruitment.

A strong recruiter creates substantial value.

The problem begins when an intermediary becomes an opaque gatekeeper, particularly when the resulting decision cannot later be reconstructed or reviewed.

A candidate submits a CV, describes experience, writes a cover letter and asks professional questions.

The candidate may then receive a message such as: "We carefully reviewed your application, but we are not ready to invite you to the next stage."

What exactly was carefully reviewed?

Which requirements were compared with the candidate's experience?

Which criterion was not met?

Who made that determination?

Was the future manager involved?

Did anyone with expertise in the field in which the candidate would actually work see the CV?

The candidate will usually not receive answers to all of these questions.

More importantly, the organization itself may not have them either.

The hiring manager sees the people whom recruitment decided to show.

The manager rarely sees the potentially strong candidates whom recruitment decided not to show.

The manager can therefore evaluate the shortlist, but has little basis for assessing the quality of the rejected-candidate pool.

A weak candidate who advances can be identified later in an interview.

A strong candidate who is rejected incorrectly may never be identified at all.

That candidate simply disappears from the organization's field of view.

An organization that wants to understand the quality of its early hiring decisions should therefore periodically conduct a blind audit of rejected applications.

The central question to the professional reviewer is simple: "If this CV had reached you directly, would you have wanted to speak with this person?"

If hiring managers or domain reviewers repeatedly answer "yes" for candidates who were originally rejected, the organization has identified a measurable level of potential false negatives in its own screening system.

The next questions become managerial rather than personal:

- How many potentially strong professionals may have been lost?
- Which criteria generate the most disagreement?
- How often does a professional reviewer reverse or materially change the original assessment?
- Where in the process do these disagreements originate?

If an organization expects an audit trail for a payment, a change to a financial model, access to data or the release of financial reporting, why should a hiring decision capable of affecting both a person's professional trajectory and the organization's access to talent sometimes be reduced to a single click: Reject?

A robust hiring process should not depend on an assumption that any individual reviewer is infallible.

It should depend on the auditability of consequential decisions.

Recruitment should not be a wall between professionals.

It should be infrastructure that helps them find one another.

21. A Call for Verification, Not Public Shaming

If you are a candidate, documenting the process is more useful than accusing a particular person without evidence.

Keep:

- the vacancy text;
- dates;
- relevant correspondence;
- candidate questions;
- responses received.

Separate what you can observe from what you infer.

If you are a hiring manager, ask periodically to see a random sample of rejected CVs - not only the shortlist.

If you are a recruiter, consider proposing an audit of your own decisions.

For a strong professional, auditability can protect reputation rather than threaten it.

If you are an HR Director, measure not only the speed of hiring but also the quality of negative decisions.

If you are a CEO, CFO, COO or functional leader, ask one question: Who currently has the authority to remove a potentially strong candidate from our hiring funnel before that candidate reaches the business - and how do we test the quality of that authority?

Public attention is valuable only when it leads to a better process.

The objective is not to create a new conflict between candidates and HR.

It is to remove unnecessary opacity between people who might otherwise have been able to work together.

The authority to make an early rejection decision should have a basis that can be checked.

Appendix A. Minimum Scorecard

No.	Field	Content / Format
1	Vacancy Criterion ID	Identifier of the vacancy criterion
2	Criterion Type	Mandatory / Trainable / Nice-to-have
3	Evidence in CV / Application	Relevant evidence contained in the CV or application
4	Evidence Status	Absent / Unclear / Confirmed

No.	Field	Content / Format
5	Recruiter Confidence	High / Medium / Low
6	Decision	Advance / Reject / Domain Review
7	Reason Code	Applicable reason code
8	Domain Review Required	Yes / No; route to Domain Review where required
9	Open Question for Candidate	Unresolved candidate question / clarification required
10	Time and Reviewer	Timestamp and Reviewer ID

For complex roles, the process may prohibit a Reject decision when the recruiter has Low confidence on a professional criterion. In such a case, the application is automatically routed to Domain Review.

Appendix B. Reason Codes

Code	Basis
R01	Qualification explicitly required by law is absent.
R02	Mandatory certification or licence is absent.
R03	A mandatory experience criterion has not been demonstrated.
R04	Location or right-to-work requirement is not met.
R05	Compensation expectations are confirmed to be outside the applicable range.
R06	Availability or required start-date condition is not met.
R07	Duplicate application or candidate withdrawal.
R08	Vacancy closed or placed on hold.
R09	Professional relevance is unclear: Domain Review required.
R10	Comparative rejection after Structured Assessment.
R11	Candidate unavailable after documented contact attempts.
R12	Other reason: mandatory written explanation required.

These codes are part of the proposed Hiring Governance control architecture. They are not presented as an established industry-standard taxonomy. R09 is especially important: uncertainty about professional relevance should be made visible and routed to Domain Review rather than silently converted into a negative professional judgment.

Appendix C. Testable Research Questions

ID	Testable Question
Q01	What proportion of early rejections are changed by a domain reviewer during blind re-review?
Q02	Which reason codes are associated with the highest Domain Override Rate?
Q03	How strongly do rejection rates vary between recruiters after controlling for vacancy mix?

ID	Testable Question
Q04	Which role requirements most often produce disagreement between the recruiter and the manager?
Q05	Does the false-negative rate decline after calibration training?
Q06	Does candidate experience improve after introducing service levels and reason categories?
Q07	How does shortlist diversity change under partially blinded initial screening?
Q08	How often do substantive candidate questions remain unanswered?

These are research questions, not conclusions.

22. Why This Problem Matters Especially for Hybrid and Non-Linear Careers

Traditional CV screening becomes especially vulnerable when a candidate does not fit neatly into a single professional category.

The labor market increasingly requires hybrid professional profiles: people who combine domain expertise, data, technology and business context.

Examples include:

- Finance + Data;
- FP&A + BI;
- Accounting + Automation;
- Product + Analytics;
- Legal + Technology.

The more interdisciplinary the profile, the less reliable simple keyword matching can become.

A candidate may not have the exact job title listed in the vacancy and may still possess a combination of capabilities that is highly relevant to the work.

This is precisely where domain review becomes important.

It is not a privilege granted to the candidate.

It is a business control against the organization's own classification error.

For experienced professionals, the problem can become more difficult because career paths are often non-linear.

A career break, a move between industries, a short employment episode, an entrepreneurial project, the study of a new technology or an independently developed project can all be interpreted in different ways.

An unstructured reviewer may treat unusual career history as a negative signal.

A domain-qualified manager can ask a different set of questions:

- What problems did the person actually solve?
- What level of responsibility did they hold?
- What artifacts did they create?
- What decisions or analytical work can they reproduce?

A mature hiring process should therefore assess evidence of capability, not merely the visual smoothness of a career chronology.

This does not mean that every unusual profile is strong.

The proposition is narrower: Unusual does not automatically mean unsuitable.

If an organization publicly emphasizes innovation, digital transformation and cross-functional work while its initial screening process systematically favors only linear career histories, there may be a contradiction between the organization's strategy and its mechanism of access.

23. Counterarguments - and Why the Governance Model Still Matters

Counterargument 1: Hiring managers cannot physically read every CV

That is true. The proposed model does not require them to. It calls for sample-based audit, escalation of ambiguous cases and a second review only where the cost of a false negative is higher. This is risk-based control, not universal re-review.

Counterargument 2: Recruiters are specifically hired to screen candidates

Also true. But the existence of a function does not establish the infallibility of its decisions. A financial controller exists to perform control, yet the controller's work may still be reviewed or audited. The more consequential the function, the more reasonable it is to measure the quality of its decisions.

Counterargument 3: Hiring managers are biased too

Absolutely. Domain review is not a magical source of truth. That is why the model does not simply replace "subjective HR" with a "subjective manager." It uses structure, independent assessments, panels, work samples and audits of disagreement. The objective is to reduce dependence on a single subjective point of judgment.

Counterargument 4: Explaining every rejection creates legal risk

That concern may be legitimate. The paper does not require employers to publish internal discussions or provide detailed personalized feedback to every rejected candidate. The requirement is narrower: the organization itself should have a professionally relevant basis for the decision. External communication may remain standardized, subject to applicable law and company policy.

Counterargument 5: Metrics can distort behavior

They can. Any KPI can be gamed. That is why FNRR or DOR should not become single-target performance measures. They should be used diagnostically, together with qualitative calibration, vacancy complexity and later process outcomes. The objective is not to optimize one number. It is to understand and improve the decision process.

24. From Personal Experience to a Public Research Agenda

The next logical step after publishing an essay like this is not to collect only stories about "bad recruiters." It is to build a structured body of observations.

Candidates could anonymously record variables such as:

- type of vacancy;
- application date;
- time to first review;
- whether a cover letter was included;
- whether the candidate asked a substantive question;
- type of response received;
- whether an interview occurred;
- whether substantive feedback was provided.

At an initial stage, this could be done without naming recruiters or companies.

The aim would be to study the process, not to create personalized campaigns against individuals.

A stronger research design would involve employers directly.

Organizations could voluntarily participate in a Random Rejection Audit and publish aggregated results.

A company willing to disclose not only time-to-hire, but also an auditable measure of the quality of its rejection decisions, could help establish a more transparent approach to hiring-process governance.

The more important precedent would not be a public list of “bad HR.”

It would be a method that other organizations could reproduce and test.

For Kazakhstan, an especially valuable next step would be an academic or independent correspondence study adapted to the local labor market and legal context.

Until such local evidence exists, strong quantitative claims about the scale of discrimination in Kazakhstan should be avoided.

But the absence of local prevalence data is not an argument for the absence of controls.

It is an argument for collecting better evidence.

Figure 1. Hiring Governance Framework

1 Vacancy Contract Must-have, trainable and nice-to-have criteria; constraints and administrative conditions Criteria defined before publication	2 Structured Intake Application ID, vacancy version and complete candidate-materials package Observability	3 Eligibility Gate Only objective constraints and pre-defined grounds Administrative filter	4 Professional Relevance Gate Rubric, evidence-based assessment and Domain Review where professional relevance is uncertain Professional relevance
5 Structured Assessment Interviews, work samples and cases against role criteria; independent ratings Comparability	6 Decision & Communication Reason code, evidence reference, decision owner and clear candidate communication Auditable decision	7 Audit & Learning Random Rejection Audit, calibration, drift monitoring and feedback Improvement loop	

CONTROL LOOP: Criteria → Evidence → Decision → Communication → Random Rejection Audit → Learning → Criteria Update

PRINCIPLE

Uncertainty about professional relevance should trigger Domain Review, not automatic rejection.

Figure 2. Recruitment Audit Trail

This figure visualizes the control chain. Its purpose is to show which elements should make it possible to reconstruct the logic of an early decision after the decision has been made.

RECRUITMENT AUDIT TRAIL

Control chain: from application to the ability to reconstruct the decision rationale

1 Application	2 Vacancy	3 Assessment	4 Decision	5 Communication	6 Audit
Application ID and CV version	Requirements version and criteria	Criterion + evidence	Decision owner + reason code	Status and response to the candidate	Sample re-review
Identifier	Context	Decision basis	Accountability	Service level	Control

KEY PRINCIPLE

Every early rejection should leave enough evidence for an authorized reviewer to reconstruct the decision rationale without guesswork.

Application → Criterion → Evidence → Decision Owner → Communication → Re-review

Figure 3. Core Management Metrics Dashboard

The dashboard does not substitute for organization-specific data. It defines a minimum set of measurements and the management meaning of each indicator.

CORE MANAGEMENT METRICS: FIRST LEVEL

Minimum diagnostic set: candidate loss, expertise calibration, auditability and process speed

FNRR False Negative Review Rate What proportion of early rejections does an independent domain reviewer move at least to the next stage? Risk of invisible candidate loss	DOR Domain Override Rate How often does a domain reviewer change the original professional assessment after independent re-review? Calibration and expertise gap
RCC Reason-Code Completeness Is there a specific reason linked to role requirements and a reference to a verifiable decision basis? Decision auditability	TFR Time-to-First-Review How much time elapses from application to the first substantive review, rather than merely opening the application record? Speed, workload and operational risk

FNRR, DOR, RCC and TFR are proposed diagnostic metrics. Thresholds should be established after baseline measurement and the pilot.

Figure 4. Hiring Governance Maturity Model

From opaque early rejection to an auditable and calibrated decision system

0 Opaque	1 Documented	2 Structured	3 Cross-Functional	4 Auditable	5 Evidence-Based
Rejection without a verifiable basis	Reason codes and basic recording	Defined criteria and a common assessment rubric	Domain Review + Four-Eyes Principle	Random Rejection Audit + diagnostic metrics	Calibration, experiments and learning
No trace	A record exists	An assessment structure exists	A second review exists	A control exists	An improvement loop exists

HOW TO READ THE RESULT

The level describes the current system, not the quality of HR as a function. The next step should be determined by the organization's own evidence.

HOW TO READ THE RESULT

The statement "Level 2 - the process is structured but not systematically audited" means that defined criteria and an assessment structure are in place, while the quality of early rejections is not yet reviewed regularly. The number of recommendations and the roadmap horizon should be derived from actual diagnostic evidence, not assigned in advance.

Figure 5. 30-Day Diagnostic Pilot

30-DAY DIAGNOSTIC PILOT

Diagnose → Test → Explain → Change → Re-test → Decide

1 Days 1-5 Baseline	2 Days 6-10 Blind Re-review	3 Days 11-15 Root-Cause Analysis	4 Days 16-22 Controlled Change	5 Days 23-27 Follow-up Sample	6 Days 28-30 Decision
Vacancies, criteria, sample, privacy	Independent domain assessments	Cause taxonomy + process analysis	Reason codes + Four- Eyes + escalation	Before/after metrics + workload	Stop / Refine / Scale
Identify bottlenecks	Observe re-review results	Build the cause map	Define the target model	Assess Evidence of Effect	Produce the Decision Memo

CONTROL LOGIC

Each phase has a question, method, evidence, deliverable and checkpoint. Progression to the next phase occurs only after the previous phase has been reviewed.

Logic: Diagnose → Test → Explain → Change → Re-test → Decide. This is the structure of a diagnostic project: each phase has a question, method, evidence, result and checkpoint.

Period	Phase	Action	Output	Checkpoint
Days 1-5	Baseline	Vacancies, criteria, sample, privacy	Pilot Charter	Scope
Days 6-10	Blind Re-review	Independent domain reviewers	Re-review Assessments	Is there disagreement?
Days 11-15	Root-Cause Analysis	Cause taxonomy + process analysis	Cause Map	Failure modes
Days 16-22	Controlled Change	Reason codes, Four- Eyes, escalation	Target Model	Ready to test
Days 23-27	Follow-up Sample	Before/after metrics	Evidence of Effect	Better / No Change / Worse
Days 28-30	Decision	Stop / Refine / Scale	Decision Memo	Launch / Stop / Refine

Why Hybrid Profiles Increase the Importance of This Framework

The labor market increasingly requires hybrid professionals - people who combine domain expertise, data, technology and business context.

Such profiles are particularly easy to misclassify when initial screening relies on narrow job titles, keywords or a linear career history.

The governance approach does not require non-standard profiles to advance automatically.

It requires the organization to distinguish a demonstrated mismatch from a situation in which professional relevance cannot be assessed reliably without domain expertise.

Final Question

FOR CEOS, HIRING MANAGERS AND HR LEADERS

FINAL QUESTION

When did we last review not the shortlist that recruitment showed the business, but a random sample of professionals whom recruitment decided not to show the business?

If the answer is “never,” the organization knows the quality of the visible part of its hiring funnel, but it does not know its own risk of invisible false-negative decisions.

References and Sources

- [1] hh.kz. “Просмотр отклика на вакансию работодателем” [Employer review of a job application]. Knowledge Base, 2026. <https://feedback.hh.kz/knowledge-base/article/kz-16191>
- [2] Bertrand, M.; Mullainathan, S. “Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination.” *American Economic Review*, 94(4), 2004, 991-1013. DOI: 10.1257/0002828042002561. <https://doi.org/10.1257/0002828042002561>
- [3] Quillian, L.; Pager, D.; Hexel, O.; Midtbøen, A. H. “Meta-analysis of field experiments shows no change in racial discrimination in hiring over time.” *Proceedings of the National Academy of Sciences*, 114(41), 2017, 10870-10875. DOI: 10.1073/pnas.1706255114. <https://doi.org/10.1073/pnas.1706255114>
- [4] Lippens, L.; Vermeiren, S.; Baert, S. “The state of hiring discrimination: A meta-analysis of (almost) all recent correspondence experiments.” *European Economic Review*, 151, 2023, 104315. DOI: 10.1016/j.euroecorev.2022.104315. <https://doi.org/10.1016/j.euroecorev.2022.104315>
- [5] Schmidt, F. L.; Hunter, J. E. “The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings.” *Psychological Bulletin*, 124(2), 1998, 262-274. DOI: 10.1037/0033-2909.124.2.262. <https://doi.org/10.1037/0033-2909.124.2.262>
- [6] Levashina, J.; Hartwell, C. J.; Morgeson, F. P.; Campion, M. A. “The Structured Employment Interview: Narrative and Quantitative Review of the Research Literature.” *Personnel Psychology*, 67, 2014, 241-293. DOI: 10.1111/peps.12052. <https://doi.org/10.1111/peps.12052>
- [7] CIPD. “Selection methods”; “Inclusive recruitment: Guide for people professionals.” Current online guidance. <https://www.cipd.org/en/knowledge/factsheets/selection-factsheet/>
- [8] Labour Code of the Republic of Kazakhstan, Articles 4, 5, 6 and 25. Adilet Legal Information System. Version current as of September 2026. <https://old.adilet.zan.kz/eng/docs/K1500000414>
- [9] gov.kz. “Недопустимость дискриминации при приеме на работу” [Inadmissibility of discrimination in hiring]. September 2026. <https://www.gov.kz/memleket/entities/abai-enbek/press/news/details/1243077?lang=ru>
- [10] U.S. Department of Justice, Civil Rights Division. “Algorithms, Artificial Intelligence, and Disability Discrimination in Hiring.” 12 May 2022. <https://www.ada.gov/resources/ai-guidance/>

NOTE

International studies are cited as evidence of general mechanisms of bias and of the limitations of unstructured selection, not as statistical estimates of the prevalence of discrimination in Kazakhstan. Separate local audit and correspondence studies are required for Kazakhstan.

Research home: serikzhanumarov.com

U Cool Research - independent research into decision quality, governance, analytics and technology.