



The Mid-Market Buyer's Guide

How to evaluate AI hiring tools when you're
not locked into an enterprise suite

2026 Edition

INTRODUCTION

Why This Guide Exists

The mid-market is in a unique position: too complex for basic tools, too agile for bloated enterprise suites. You're running 6-10 specialized HR tools, hiring across multiple roles and geographies, and now AI vendors are flooding your inbox.

How do you separate substance from sales pitch?

The Problem

- **78% of companies already deploy AI in hiring** -- many mid-market firms are adopting without a framework for evaluating what they're buying
- **Enterprise players** (Workday, SAP, Oracle) bundle AI into suites you don't use; point solutions promise the moon with no proof
- **Regulatory pressure is accelerating:** NYC LL144 is active now, EU AI Act hits August 2026, Colorado and California close behind
- **Only 26% of job applicants trust AI** will evaluate them fairly (Gartner) -- your candidates are watching

What This Guide Covers

A practical, vendor-neutral framework for evaluating AI hiring tools -- built specifically for mid-market companies that need flexibility, compliance, and real ROI without a six-figure consulting engagement to figure it out.

Inside you'll find:

- A **five-pillar evaluation framework** with scoring rubrics
- A printable **vendor comparison scorecard** (24 criteria, scored 1-5)
- **20 questions** to ask during vendor demos
- **Red flags and green lights** for instant vendor screening
- An **ROI calculator worksheet** to build your business case
- A **30-day evaluation timeline** from research to decision

The Mid-Market Advantage (and Blind Spots)

Why Mid-Market Companies Have the Upper Hand

- **Best-of-breed flexibility.** Unlike enterprise (locked into legacy suites) or SMB (price-constrained), mid-market can pick the best tool for each job.
- **Faster implementation cycles.** No 18-month enterprise rollouts -- the right tools go live in days or weeks.
- **Direct access to decision-makers.** HR leaders can evaluate, pilot, and deploy without five layers of procurement.
- **Budget capacity for premium tools.** Mid-market companies (\$10M-\$1B revenue, 100-1,000 employees) typically invest \$100K-\$500K/year in HR technology -- enough for enterprise-grade solutions without enterprise complexity.

The Blind Spots to Watch

- **Compliance without a compliance team.** You don't have a Chief AI Ethics Officer, but regulators don't care -- liability is the same. NYC Local Law 144 penalties run \$500-\$1,500 per violation per candidate. For 1,000 candidates, that's up to \$1.5M in exposure.
- **Vendor lock-in risk.** Some "AI" tools are really data traps -- your candidate data goes in but never comes back out. Always ask: what happens to my data if I leave?
- **The "free trial" trap.** Tools that work at 50 hires/year collapse at 500 -- scalability matters. Ask for performance benchmarks at your actual hiring volume.
- **Integration debt.** A tool that doesn't connect to your ATS/HRIS stack creates more work, not less. If you're running Greenhouse plus BambooHR, the tool needs to slot in -- not replace everything.
- **The trust gap.** 78% of companies deploy AI in hiring, but only 31% have enforcement-level governance. That 47% gap is where the lawsuits land. Your candidates, your legal team, and now regulators all want to know: can you explain how your AI makes decisions?

Key insight: Mid-market companies represent the highest-ROI segment for AI hiring tools. Enterprise is locked into suites. SMB can't afford premium solutions. You can move fast, choose the best, and implement in days -- if you know what to look for.

The Five-Pillar Evaluation Framework

A structured approach for comparing AI hiring vendors. Score each pillar during demos and reference calls to build an objective, apples-to-apples comparison.

Pillar 1: Scientific Validity

Does the tool actually predict job performance -- or just automate resume sorting?

What to ask:

- What peer-reviewed research underpins your methodology?
- What is the standalone predictive validity ($r = ?$) for job performance?
- How does it perform when combined with structured interviews?
- Are role-specific profiles built from YOUR top performers -- or generic industry models?
- Can the assessment be gamed with preparation or AI-generated responses?

Rating	Criteria
Strong	Rooted in peer-reviewed I/O psychology research; $r \geq .35$ standalone; role-specific profiling from your company data
Acceptable	Published validity data; some customization available; moderate predictive power
Weak	No published research; "proprietary" methodology with no transparency; keyword matching only

Benchmark: 85 years of I/O psychology meta-analysis (Schmidt & Hunter, 1998) establishes cognitive ability + structured interviews ($r = .63$) as the gold standard for predicting job success. Any tool claiming higher validity without peer-reviewed evidence should be questioned.

Pillar 2: Compliance & Audit-Readiness

Will this tool protect you from regulatory risk -- or create it?

What to ask:

- Is bias testing performed BEFORE deployment or only after complaints?
- Does the tool generate documentation for NYC LL144, EU AI Act, Colorado AI Act, and California's 4-year retention mandate?
- Does the vendor support independent third-party bias audits?
- Who assumes liability if the tool produces discriminatory outcomes -- you or the vendor?
- How frequently is adverse impact testing conducted, and across which demographic groups?

Rating	Criteria
Strong	Pre-deployment bias validation; automatic audit documentation; supports third-party audits; continuous monitoring
Acceptable	Annual bias audits; manual compliance documentation; responsive to audit requests
Weak	No bias testing; no documentation; "we're compliant" with no evidence; you assume all risk

Reality check: The EEOC has stated that employers remain liable for discriminatory AI outcomes even if the tool is designed by a third-party vendor. "We're compliant" without specifics is a red flag, not a reassurance.

Pillar 3: Transparency & Explainability

Can you (and your candidates) understand how decisions are made?

What to ask:

- Are scoring decisions explainable in plain language -- not just a number?
- Can a recruiter explain to a candidate WHY they scored the way they did?
- Is the model a "black box" or does it provide decision pathways?
- How is AI use disclosed to candidates, and at what stage?
- Do candidates have the option to request an alternative assessment method?

Rating	Criteria
Strong	Plain-language score narratives; full decision audit trails; proactive candidate disclosure; alternative assessment available
Acceptable	Summary-level explanations; audit trails available on request; candidate notification provided
Weak	No explanations; "proprietary algorithm" justification; no candidate disclosure process

Trust data: 79% of candidates want to know exactly how AI is being used in their hiring process (Josh Bersin). Transparency isn't just ethical -- it's a competitive advantage for candidate experience and employer brand.

Pillar 4: Integration & Implementation

Will it work with what you already have -- and how fast?

What to ask:

- Which ATS and HRIS systems do you integrate with natively?

- Is the architecture API-first, or does integration require custom development?
- What is the realistic implementation timeline (not the sales deck version)?
- What does your data model look like -- and can I export my data if I leave?
- What training and onboarding support is included?

Rating	Criteria
Strong	API-first; native integrations with major ATS/HRIS; live in 2 weeks or less; data portability guaranteed; dedicated implementation specialist
Acceptable	API available; most common integrations; live in 1-2 months; standard onboarding
Weak	Manual data transfer; limited integrations; 3+ month implementation; self-service only

Mid-market reality: You're running Greenhouse or Lever plus BambooHR or ADP -- not a monolithic SAP environment. The tool needs to slot into your stack, not replace it. Ask for a live demo of the integration, not a slide.

Pillar 5: ROI & Total Cost of Ownership

What will this actually cost -- and what will you get back?

What to ask:

- What is the full TCO including implementation, training, ongoing support, and compliance maintenance?
- How do you measure ROI -- time-to-hire alone, or correlated to job performance and retention?
- What retention improvements can you demonstrate (with customer data, not projections)?
- Are there per-hire fees, volume tiers, or hidden costs?
- What does year-two pricing look like?

Rating	Criteria
Strong	Transparent all-in pricing; ROI tied to retention + performance outcomes; customer case studies with data; no hidden fees
Acceptable	Clear pricing model; ROI measured by efficiency gains; reference customers available
Weak	Opaque pricing; ROI claims with no supporting data; per-hire fees that scale unpredictably

TCO cost buckets to request from every vendor:

1. Software licensing (annual subscription)
2. Implementation and onboarding
3. Integration development (if any)

4. Training and change management
5. Ongoing compliance and bias audit support
6. Year-over-year price escalation

REGULATORY LANDSCAPE

The Compliance Landscape at a Glance

AI hiring regulation is accelerating. Here's what mid-market companies need to know about the six frameworks that affect your vendor selection.

Regulation	Status	Key Requirement	Penalty Exposure	Mid-Market Impact
NYC Local Law 144	Active NOW	Annual bias audit + public disclosure	\$500-\$1,500 per violation per candidate	Any company hiring in NYC
EU AI Act	Aug 2026	High-risk classification; conformity assessment	Up to EUR 35M or 7% global revenue	Any company with EU candidates
Colorado AI Act	Jun 2026	Annual impact assessments for high-risk AI	Enforcement TBD	CO-based or hiring in CO
California FEHA AI	Active	4-year data retention; bias audit evidence	DFEH investigation + penalties	CA-based or hiring in CA
Illinois HB3733	Jan 2026	No proxy variables; employee notification	IHRA enforcement	IL-based or hiring in IL
EEOC Title VII	Ongoing	Four-fifths rule; employer liability for vendor AI	Class action: typically \$1M+	Every US employer

Cross-reference: For a detailed 24-item compliance self-assessment with checkbox scoring, see the companion *AI Compliance Checklist* available at catalyze.com/resources.

What This Means for Your Vendor Evaluation

Every AI hiring tool you evaluate should be able to demonstrate compliance readiness for the jurisdictions where you hire. Don't accept vague assurances. Ask for specific documentation samples and audit procedures for each regulation that applies to your operations.

If a vendor can't name the specific regulations they support -- or can't show you a sample audit report -- that tells you everything you need to know about their compliance maturity.

SCORECARD

Vendor Evaluation Scorecard

Score each vendor 1-5 on every criterion during demos and reference calls. Use this as a printable comparison worksheet.

Evaluation Criteria	Vendor A	Vendor B	Vendor C
SCIENTIFIC VALIDITY			
Peer-reviewed research foundation	___ / 5	___ / 5	___ / 5
Published predictive validity data (r value)	___ / 5	___ / 5	___ / 5
Role-specific profiling (vs. generic)	___ / 5	___ / 5	___ / 5
Assessment resistant to gaming / AI manipulation	___ / 5	___ / 5	___ / 5
Combined validity with structured interviews	___ / 5	___ / 5	___ / 5
COMPLIANCE & AUDIT-READINESS			
Pre-deployment bias testing	___ / 5	___ / 5	___ / 5
Automatic audit documentation generation	___ / 5	___ / 5	___ / 5
Multi-jurisdiction compliance support	___ / 5	___ / 5	___ / 5
Third-party audit support	___ / 5	___ / 5	___ / 5
Continuous adverse impact monitoring	___ / 5	___ / 5	___ / 5
TRANSPARENCY & EXPLAINABILITY			
Plain-language score explanations	___ / 5	___ / 5	___ / 5
Full decision audit trails	___ / 5	___ / 5	___ / 5
Candidate disclosure process	___ / 5	___ / 5	___ / 5
Human-in-the-loop (AI informs, not decides)	___ / 5	___ / 5	___ / 5
INTEGRATION & IMPLEMENTATION			
Native ATS/HRIS integrations	___ / 5	___ / 5	___ / 5
API-first architecture	___ / 5	___ / 5	___ / 5
Implementation timeline (faster = higher)	___ / 5	___ / 5	___ / 5
Data portability / no lock-in	___ / 5	___ / 5	___ / 5
Onboarding and training quality	___ / 5	___ / 5	___ / 5
ROI & TOTAL COST OF OWNERSHIP			

Evaluation Criteria	Vendor A	Vendor B	Vendor C
Transparent, all-in pricing			
ROI tied to retention + performance	___ / 5	___ / 5	___ / 5
Customer case studies with data	___ / 5	___ / 5	___ / 5
Year-over-year pricing clarity	___ / 5	___ / 5	___ / 5
TOTAL SCORE	___ / 120	___ / 120	___ / 120

Scoring Interpretation

- **96-120:** Strong candidate -- move to pilot
- **72-95:** Promising -- dig deeper on weak areas
- **48-71:** Gaps exist -- require vendor remediation plan before proceeding
- **Below 48:** Significant risk -- reconsider

DEMO PREPARATION

20 Questions to Ask During Your Vendor Demo

Organized by pillar, designed to separate substance from sales pitch. Print this page and bring it to every demo.

Scientific Validity (Questions 1-4)

1. What peer-reviewed research is your assessment methodology built on? Can you share the citations?
2. What is your tool's predictive validity coefficient for actual job performance (not just hiring speed)?
3. How do you build role-specific profiles -- from our company's top performers or from generic industry data?
4. How resistant is the assessment to gaming by candidates using AI tools to prepare?

Compliance & Risk (Questions 5-8)

5. Walk me through your bias testing process -- when does it happen relative to deployment?
6. Which specific regulations does your tool generate compliance documentation for? Show me an example.
7. If your tool produces a discriminatory outcome, what is your contractual liability position?
8. How do you handle intersectional bias testing (e.g., race x gender, age x ethnicity)?

Transparency (Questions 9-12)

9. Show me what a candidate's score explanation looks like -- can a recruiter explain it to a hiring manager in plain language?
10. Does the tool make hiring decisions, or does it inform human decisions? Where is the human checkpoint?

-
11. What do candidates see and experience? Can I take the assessment myself before we buy?
 12. How and when are candidates notified that AI is being used in their evaluation?

Integration & Implementation (Questions 13-16)

13. Show me a live integration with [our ATS] -- not a slide, the actual data flow.
14. What is your actual median implementation timeline for companies our size? Can I talk to a reference?
15. If we leave, how do we get our data out? In what format, and within what timeframe?
16. What does your customer support model look like post-implementation -- dedicated CSM or ticket queue?

ROI & Cost (Questions 17-20)

17. Give me your all-in cost for year one AND year two -- no surprises.
18. Show me a customer case study where you can prove retention improvement, not just faster time-to-hire.
19. Are there per-hire fees, assessment volume caps, or overage charges?
20. What training is included, and what costs extra?

QUICK SCREENING

Red Flags and Green Lights

Use this as a quick-reference during initial vendor screening. These signals help you shortlist fast before investing time in full demos.

Red Flags -- Walk Away If:

- ✗ "Our algorithm is proprietary" with zero published research or validity data
- ✗ No bias testing performed before deployment -- only reactive after complaints
- ✗ Vendor won't share their contractual liability position on discriminatory outcomes
- ✗ Implementation requires 3+ months and dedicated IT resources from your team
- ✗ No data portability -- your candidate data is locked in their system
- ✗ "We're compliant" with no specific regulation references or documentation samples
- ✗ Pricing requires a sales call with no published ranges or model transparency

Green Lights -- Promising Signals:

- ✓ Methodology grounded in peer-reviewed I/O psychology with published validity coefficients
- ✓ Pre-deployment bias validation with continuous monitoring post-launch
- ✓ Automatic compliance documentation for specific jurisdictions (not generic "we handle compliance")
- ✓ API-first architecture with native integrations for your existing ATS/HRIS stack

- ✓ Plain-language score explanations that recruiters and candidates can understand
- ✓ Transparent pricing with clear year-over-year terms and no hidden fees
- ✓ Customer references willing to share retention and performance outcome data
- ✓ Human-in-the-loop design -- AI informs decisions, humans make them

BUSINESS CASE

ROI Calculator Worksheet

Build your business case with this three-step worksheet. Use conservative estimates to start -- the actual numbers will likely be higher.

Step 1: Calculate Your Current Cost of Bad Hires

Input	Your Number
Annual hires	_____
Average cost per hire (recruiting + onboarding + training)	\$_____
Current turnover rate (within first year)	_____ %
Annual cost of turnover = (Hires x Cost x Turnover Rate)	\$_____

Step 2: Project AI-Driven Improvement

Scenario	Conservative (15%)	Moderate (25%)	Strong (30%)
Turnover reduction	_____ %	_____ %	_____ %
Annual savings	\$_____	\$_____	\$_____

Step 3: Calculate Net ROI

Item	Amount
Annual savings from reduced turnover	\$_____
Additional savings: reduced time-to-hire (recruiter hours)	\$_____
Additional savings: compliance risk mitigation	\$_____
Total annual benefit	\$_____
Minus: Year-one TCO (software + implementation + training)	(\$_____)

Item	Amount
Net Year-One ROI	\$ _____

Example: 500 annual hires x \$5,000 cost per hire x 40% turnover = \$1M current cost. A 30% turnover reduction saves \$300K/year. If the tool costs \$100K all-in, that's a 3:1 ROI in year one.

IMPLEMENTATION

Implementation Roadmap

What a good implementation looks like. Use this as a benchmark when evaluating vendor timelines.

Phase	Timeline	Activities	Your Role
1. Connect	Day 1	API credentials exchanged; ATS/HRIS connection established and verified	Provide system access + point of contact
2. Configure	Day 2-3	Assessment models calibrated to your roles, requirements, and compliance needs	Share top-performer profiles + role requirements
3. Validate	Day 4-5	Test with real candidate data; review results with your team; fine-tune scoring	Participate in review sessions; provide feedback
4. Go Live	Day 5-7	Launch to all users; ongoing support and optimization from dedicated specialist	Internal announcement; recruiter training

Warning sign: If a vendor quotes 3+ months for implementation, they're either building custom integrations (costly) or lack API-first architecture (inflexible). Mid-market implementations should be measured in days, not quarters.

DECISION TIMELINE

Your 30-Day Evaluation Process

From first research to signed contract in four weeks. Here's the timeline.

Week 1: Research & Shortlist

- Complete the vendor scorecard (pages 7-8) for 3-4 candidates
- Eliminate any with red flags from page 11
- Schedule demos with top 2-3 vendors

Week 2: Deep-Dive Demos

- Use the 20 questions (pages 9-10) during each demo

-
- Request to take the assessment yourself -- experience what candidates will see
 - Ask for 2 customer references at companies your size

Week 3: Reference Checks & Pilot Design

- Call references -- ask specifically about retention outcomes, implementation reality, and support quality
- Negotiate a pilot scope: specific role, defined success metrics, 30-60 day timeline

Week 4: Decision & Negotiate

- Compare final scorecard totals across all vendors
- Negotiate contract terms: data portability, liability allocation, pricing lock, exit clause
- Confirm implementation timeline and dedicated support resources

ABOUT CATALYZR

Built for the Mid-Market Moment

Catalyzr's CQ Intelligence Engine was purpose-built for companies that need enterprise-grade predictive hiring without the enterprise overhead.

Why mid-market teams choose Catalyzr:

- **Scientific foundation.** Built on 85 years of peer-reviewed I/O psychology research. CQ assessment predictive validity: $r = .40+$ standalone, $r = .63$ combined with structured interviews.
- **Compliance-first.** Pre-deployment bias validation with automatic documentation for NYC LL144, EU AI Act, Colorado, California, and EEOC. Audit-ready from day one.
- **API-first integration.** Connects to Greenhouse, Lever, iCIMS, ADP, BambooHR, Paylocity, and SAP SuccessFactors. Live in 5-7 days with a dedicated implementation specialist.
- **100% explainable.** Every CQ score comes with a plain-language narrative explaining the match -- no black boxes. Recruiters can explain results to candidates and hiring managers.
- **SAP SuccessFactors Partner.** Enterprise-grade integration for companies that need it, without the enterprise complexity.

Ready to evaluate? Book a demo at catalyzr.com/demo

Download the companion checklist: AI Compliance Checklist at catalyzr.com/resources

catalyzr.com | hello@catalyzr.com