AI Risk Assessment: TalentScreen Pro

Vesta Mutual Insurance AS · Assessment version 1.0 · 29 June 2026 · Assessor: AI Governance Lead · Approver: COO · Next review: December 2026

Worked example for portfolio and training purposes. Vesta Mutual Insurance AS is a fictional company; all data, metrics and names are invented. Prepared by Erik Bernath, Furioso AI Consulting OÜ (furiosoaiconsulting.eu), June 2026. Licensed CC BY 4.0: reuse freely with attribution. This document is informational and is not legal advice.

1. System and context

TalentScreen Pro is a procured SaaS tool that parses applications for Vesta's roughly 35 hires per year (600 to 900 applications), scores candidates against requisition profiles, and produces a ranked shortlist for recruiters. The vendor is the AI Act provider; Vesta is the deployer. The system touches job applicants (the affected persons), 4 recruiters and hiring managers (the users), and the works council (information rights under Art. 26(7)).

Classification. High-risk under Annex III 4(a) (recruitment and candidate filtering). Deployer obligations under Art. 26 apply in full from 2 August 2026 for systems of this type. A fundamental rights impact assessment under Art. 27 is not triggered for a private employer's recruitment use; the reasoning is recorded because the same question for PriceWise (insurance pricing, Annex III 5(c)) lands the other way.

2. Method

Risks were identified in a 2-hour structured workshop (HR, IT, DPO, one works-council member) using the system's data flow as the walking skeleton, then scored on a 3-level likelihood and severity scale. Severity reflects harm to candidates first, legal exposure second, operations third. The register below shows inherent risk (no controls), the controls Vesta operates, and residual risk. Scoring is judgment, recorded so it can be challenged; that is the point of writing it down.

3. Risk register

Risk	Inherent (L × S)	Inherent level	Key mitigations	Residual level
Proxy discrimination: model ranks down candidates via variables correlated with protected attributes (career gaps, school names, address)	Likely × Major	High	Quarterly disparate-impact testing across age, gender and (where inferable) ethnicity proxies on shortlist-rate data; selection-rate ratios below 0.8 trigger investigation; vendor bias-audit report required annually	Medium
Automation bias: recruiters treat the ranking as a decision and never look below the cut line	Likely × Major	High	Tool configured to shortlist, never reject; weekly sample review of 10 auto-deprioritized CVs per requisition; recruiter training includes a seeded-CV exercise; override rate tracked as a health metric	Medium
Opaque criteria: nobody at Vesta can explain a ranking to a candidate or court	Possible × Major	High	Vendor contract requires feature-importance documentation per requisition profile; HR keeps a plain-language explanation template; unexplainable rankings are not used	Medium
Parsing failure: non-standard CV formats (common for non-native applicants) score artificially low	Likely × Moderate	Medium	Structured application form as primary input, CV as secondary; quarterly test with synthetic non-standard CVs	Low
GDPR Art. 22: shortlisting amounts to automated decision-making without safeguards	Possible × Major	High	Human review of every rejection before notice; candidate privacy notice names the tool, logic and rights; DPIA completed and linked	Low-Medium

Unannounced vendor model updates change behaviour mid-campaign	Possible × Moderate	Medium	Contract clause: 30-day advance notice of model changes plus changelog; re-run of disparate-impact tests after each update	Low
Worker-information duty missed (Art. 26(7))	Possible × Moderate	Medium	Works-council briefing held and minuted before go-live; standing agenda item at annual review	Low

4. Deployer-obligation checklist (Art. 26)

- Use per vendor's instructions: instructions reviewed, deviations none. Compliant.
- Competent human oversight: 2 trained recruiters per requisition; training recorded with literacy evidence (links to AI Policy section 11). Compliant.
- Input-data relevance and quality: structured form primary; quarterly parsing tests. Compliant with monitoring.
- Monitoring and provider feedback: anomalies reported to vendor; override rate reviewed monthly. Compliant.
- Log retention: automatically generated logs retained 12 months (above the 6-month minimum). Compliant.
- Worker information (Art. 26(7)): works council briefed 14 May 2026, minuted. Compliant.
- Candidate transparency and GDPR: privacy notice updated; every rejection passes human review before notice; DPIA ref. DP-2026-04. Compliant.

5. Residual risk statement and acceptance

With controls operating, the highest residual risks are proxy discrimination and automation bias, both Medium. They are inherent to ranking tools of this class and are held at Medium by testing and sampling controls whose evidence is auditable (test results, sample logs, override metrics). The COO accepts this residual level for continued use, conditional on: no red disparate-impact result older than 30 days unresolved, override-rate review monthly, and full reassessment at any vendor model change or by December 2026.

6. Framework mapping

EU AI Act: implements deployer duties (Art. 26) and documents the Art. 27 non-applicability call. NIST AI RMF: MAP 5 (impact identification), MEASURE 2 (bias evaluation), MANAGE 1-2 (risk treatment and residual acceptance). ISO/IEC 42001: feeds the AI impact assessment requirement (Annex A) and management review input. GDPR: DPIA cross-referenced rather than duplicated.