

AI Risk Assessment Scorecard

Related Card: Plan Model Risk Mitigation | Format: 1-page PDF scorecard

Feature Name:

Assessed by:





Date:

What it does:

Risk Scoring Matrix

Instructions: Score each risk 1-5 (1 = low risk, 5 = critical risk). Multiply Impact × Likelihood = Risk Score.

| Risk Category | Impact (1-5) | Likelihood (1-5) | Risk Score | Priority | Top Mitigation |
|---|--------------|------------------|------------|----------|----------------|
| Bias Risk Unfair outcomes | | | | | |
| Privacy Risk Data exposure | | | | | |
| Performance Risk AI failures | | | | | |
| UX/Trust Risk User trust erosion | | | | | |
| Safety Risk Potential harm | | | | | |
| Compliance Risk Regulatory violations | | | | | |

- Risk Priority Key:**
-  **Critical (20-25):** Block launch until mitigated
 -  **High (12-19):** Must address before launch
 -  **Medium (6-11):** Address if time permits
 -  **Low (1-5):** Monitor post-launch

1. Bias Risk Assessment

- ☐ Training data represents all user demographics
- ☐ Tested outputs across different user groups
- ☐ Feedback mechanism for users to report bias

Key risks:

2. Privacy Risk Assessment

- ☐ Data anonymized before training
- ☐ Tested for data leakage in outputs
- ☐ User consent obtained for data usage

Key risks:

3. Performance Risk

- ☐ Acceptable error rate defined: _____%
- ☐ Tested edge cases & adversarial inputs
- ☐ Fallback exists when AI fails

Key risks:

4. UX/Trust Risk

- ☐ Users understand what AI is doing
- ☐ Users can verify/override AI decisions
- ☐ AI transparent about confidence levels

Key risks:

5. Safety Risk

- ☐ Worst-case failure scenario documented
- ☐ Safeguards against harmful outputs
- ☐ Human oversight for high-stakes decisions

Key risks:

6. Compliance Risk

- ☐ Consulted legal/compliance teams
- ☐ Can explain AI decisions if audited
- ☐ Data retention policies followed

Key risks:

Overall Risk Assessment

Critical Risks (🔴):

High Risks (🟡): _____

Medium Risks (🟡):

Low Risks (🟢): _____

Launch Recommendation

- ☐ ☒ **Green Light:** All risks mitigated or acceptable
- ☐ **⚠️ Yellow Light:** Launch with mitigations in place, monitor closely
- ☐ **🛑 Red Light:** Do not launch until critical risks addressed

Justification:

Post-Launch Monitoring

Top 3 Metrics to Track:

Review Cadence:
☐ Weekly
☐ Biweekly
☐ Monthly

Review Owner:
