

InfraRader AI

AI MVP Success Metrics & HITL Plan

AI MVP Success Metrics & HITL Plan

Comprehensive Framework for AI Model Performance and Human Validation

Confidential & Proprietary

Date: October 21, 2025

Contents

1	Executive Summary	2
1.1	Key Objectives	2
2	AI MVP Success Metrics	2
2.1	Data Quality Metrics	2
2.2	AI Model Performance Metrics	3
2.3	Business Impact Metrics	3
3	Human-in-the-Loop (HITL) Plan	4
3.1	HITL Framework Overview	4
3.2	HITL Roles and Responsibilities	4
3.3	HITL Workflow	4
3.4	HITL Implementation Phases	5
4	Quality Assurance Framework	5
4.1	Quality Gates	5
4.2	Quality Metrics Dashboard	6
5	Continuous Improvement Process	6
5.1	Model Performance Monitoring	6
5.2	Feedback Integration	7
5.3	Model Evolution	7
6	Risk Management	7
6.1	AI Model Risks	7
6.2	Data Quality Risks	7
6.3	Operational Risks	7
6.4	Compliance Risks	7
7	Success Criteria	7
7.1	MVP Success Criteria	7
7.2	Long-term Success Criteria	8
8	Conclusion	8

1 Executive Summary

This document defines InfraRader AI's AI MVP success metrics and Human-in-the-Loop (HITL) plan, providing a comprehensive framework for measuring AI model performance, ensuring data quality, and implementing continuous improvement processes.

1.1 Key Objectives

- Define measurable success metrics for AI models
- Implement Human-in-the-Loop validation processes
- Establish continuous learning and improvement cycles
- Ensure data quality and model reliability
- Create scalable AI operations framework

2 AI MVP Success Metrics

2.1 Data Quality Metrics

1. Accuracy Rate

- Target: 95% accuracy for critical project data
- Measurement: Manual verification of AI-extracted data
- Frequency: Weekly validation on sample data
- Threshold: Minimum 90% accuracy for production deployment

2. Completeness Score

- Target: 98% completeness for required fields
- Measurement: Percentage of populated fields per project
- Frequency: Daily monitoring and reporting
- Threshold: Minimum 95% completeness for active projects

3. Freshness Index

- Target: Data updates within 24-48 hours
- Measurement: Average time from source update to platform update
- Frequency: Real-time monitoring
- Threshold: Maximum 72 hours for critical updates

4. Consistency Score

- Target: 95% consistency across data sources
- Measurement: Cross-validation of data points
- Frequency: Weekly analysis
- Threshold: Minimum 90% consistency for production

2.2 AI Model Performance Metrics

1. LLM Performance

- Document extraction accuracy: 95%
- Entity recognition precision: 90%
- Entity recognition recall: 85%
- Processing speed: <5 seconds per document

2. Computer Vision Performance

- Change detection accuracy: 90%
- Progress classification accuracy: 85%
- False positive rate: <5%
- Processing speed: <30 seconds per image

3. Confidence Scoring

- Confidence score accuracy: 90%
- Calibration error: <0.1
- Coverage: 100% of data points
- Update frequency: Real-time

2.3 Business Impact Metrics

1. Customer Success

- User adoption rate: 80% within 30 days
- Feature utilization: 70% of available features
- Customer satisfaction: NPS >50
- Time to value: <2 weeks

2. Operational Efficiency

- Data processing volume: 10,000+ projects/month
- Manual review reduction: 70%
- Error rate reduction: 60%
- Processing cost reduction: 50%

3. Revenue Impact

- Customer acquisition: 10+ customers in Year 1
- Revenue growth: \$500K ARR by Month 12
- Customer retention: 90% annual retention
- Expansion revenue: 30% of total revenue

3 Human-in-the-Loop (HITL) Plan

3.1 HITL Framework Overview

The HITL system integrates human expertise with AI automation to ensure data quality, model improvement, and continuous learning.

3.2 HITL Roles and Responsibilities

1. Data Stewards

- Verify and validate AI-extracted data
- Flag inconsistencies and errors
- Provide domain expertise and context
- Maintain data quality standards

2. AI Specialists

- Monitor model performance and accuracy
- Identify improvement opportunities
- Implement model updates and refinements
- Analyze failure cases and edge cases

3. Domain Experts

- Provide industry-specific knowledge
- Validate complex project assessments
- Review risk analysis and predictions
- Ensure regulatory compliance

4. Quality Assurance

- Conduct systematic quality reviews
- Implement testing protocols
- Monitor compliance and standards
- Report on quality metrics and trends

3.3 HITL Workflow

1. Data Ingestion

- AI processes incoming data automatically
- Low-confidence data flagged for human review
- High-confidence data validated by sampling
- Edge cases escalated to domain experts

2. Quality Validation

- Systematic review of AI outputs

- Cross-validation with multiple sources
- Human verification of critical data points
- Feedback integration for model improvement

3. Model Improvement

- Analysis of human feedback and corrections
- Identification of model weaknesses
- Implementation of improvements and updates
- Validation of enhanced model performance

4. Continuous Learning

- Regular model retraining with new data
- Incorporation of human feedback
- Performance monitoring and optimization
- Knowledge base updates and expansion

3.4 HITL Implementation Phases

1. Phase 1: Foundation (Months 1-3)

- Establish HITL team and processes
- Implement basic validation workflows
- Set up quality monitoring systems
- Train team on AI systems and tools

2. Phase 2: Optimization (Months 4-6)

- Refine validation processes and criteria
- Implement advanced quality metrics
- Develop automated feedback systems
- Optimize human-AI collaboration

3. Phase 3: Scale (Months 7-12)

- Scale HITL processes for growth
- Implement advanced AI capabilities
- Develop specialized expertise areas
- Create self-improving systems

4 Quality Assurance Framework

4.1 Quality Gates

1. Data Ingestion Gate

- Source validation and verification

- Format and structure compliance
- Completeness and accuracy checks
- Security and privacy validation

2. AI Processing Gate

- Model performance validation
- Output quality assessment
- Confidence score verification
- Error detection and handling

3. Human Review Gate

- Expert validation and verification
- Cross-source consistency checks
- Domain-specific accuracy review
- Final quality approval

4. Delivery Gate

- Customer-facing quality standards
- Performance and reliability checks
- User experience validation
- Final delivery approval

4.2 Quality Metrics Dashboard

- **Real-time Monitoring:** Live quality metrics and alerts
- **Performance Tracking:** Historical trends and patterns
- **Exception Reporting:** Quality issues and resolutions
- **Improvement Tracking:** Progress on quality initiatives

5 Continuous Improvement Process

5.1 Model Performance Monitoring

- **Automated Monitoring:** Real-time performance tracking
- **Alert Systems:** Immediate notification of issues
- **Performance Analysis:** Regular review and assessment
- **Improvement Planning:** Systematic enhancement planning

5.2 Feedback Integration

- **Customer Feedback:** Direct user input and suggestions
- **Internal Feedback:** Team observations and insights
- **System Feedback:** Automated performance data
- **External Feedback:** Industry expert input

5.3 Model Evolution

- **Regular Retraining:** Scheduled model updates
- **Incremental Improvements:** Continuous small enhancements
- **Major Upgrades:** Significant capability additions
- **Version Control:** Systematic model versioning

6 Risk Management

6.1 AI Model Risks

- **Risk:** Model performance degradation
- **Mitigation:** Continuous monitoring and HITL validation

6.2 Data Quality Risks

- **Risk:** Inaccurate or incomplete data
- **Mitigation:** Multi-source validation and human review

6.3 Operational Risks

- **Risk:** HITL resource constraints
- **Mitigation:** Scalable processes and automation

6.4 Compliance Risks

- **Risk:** Regulatory or legal issues
- **Mitigation:** Comprehensive compliance framework

7 Success Criteria

7.1 MVP Success Criteria

- 95% accuracy rate for critical data extraction
- 90% user satisfaction score
- 70% reduction in manual data processing
- 10+ paying customers within 12 months

7.2 Long-term Success Criteria

- 98% accuracy rate across all data types
- 95% customer retention rate
- 90% automation of data processing
- \$5M+ ARR by Year 3

8 Conclusion

This AI MVP success metrics and HITL plan provides a comprehensive framework for ensuring AI model performance, data quality, and continuous improvement. Through systematic implementation of these processes, InfraRader AI will deliver reliable, accurate, and valuable intelligence to our customers while building a sustainable competitive advantage.