

CIRF Research Detailed Milestones & Deliverables

Phase 1 Milestones (Months 1-3)

Month 1 Deliverables

☐ **Technical Infrastructure** (Week 1-2)

- Python environment with required libraries (pandas, sklearn, spacy, selenium)
- SQLite/PostgreSQL database with complete schema
- Basic web scraping framework operational
- Data quality assessment rubric finalized

☐ **CIRF Scoring Framework** (Week 3-4)

- Detailed scoring rubric for all 13 components (0-1 scale)
- Evidence requirements for each score level
- Inter-rater reliability testing protocol
- Manual scoring templates created

Month 2 Deliverables

☐ **Pilot Dataset Collection**

- 50 documented cultural entrepreneurship cases
- Equal distribution of success/failure outcomes
- Geographic diversity across 3+ regions
- Sector diversity across 5+ categories
- Source documentation and citations complete

☐ **Data Quality Assessment**

- Quality scores assigned to all pilot cases
- Source reliability evaluation complete
- Missing data patterns identified
- Data collection efficiency metrics

Month 3 Deliverables

☐ **Methodology Validation**

- Manual CIRF scoring completed for all 50 cases
- Inter-rater reliability coefficient ≥ 0.80
- Preliminary statistical analysis complete

- Automated vs manual scoring comparison
- **Critical Decision Point:** Methodology approval for scaling

Phase 2 Milestones (Months 4-6)

Month 4 Deliverables

☐ Scaled Data Collection

- 150 additional cases collected (200 total)
- Geographic representation across 10+ countries
- Sector balance achieved across target categories
- Automated collection workflows operational

Month 5 Deliverables

☐ NLP System Development

- Component-specific NLP models trained and tested
- Confidence scoring algorithm validated
- Evidence extraction system operational
- Automated CIRF scoring achieving $\geq 75\%$ accuracy vs manual

Month 6 Deliverables

☐ Mid-Scale Analysis

- Complete statistical analysis of 200-case dataset
- Preliminary hypothesis testing results
- Predictive model baseline established
- **Critical Decision Point:** Statistical approach validation

Phase 3 Milestones (Months 7-12)

Months 7-9 Deliverables

☐ Production Data Collection

- 400 additional cases collected (600 total)
- Full automation of collection pipeline
- Quality control processes validated
- Geographic/sector targets achieved

Months 10-12 Deliverables

☐ **Advanced Analytics Framework**

- Machine learning models developed and tested
- Cross-validation protocols implemented
- Preliminary dissertation results generated
- **Critical Decision Point:** Sufficient statistical power achieved

Phase 4 Milestones (Months 13-18)

Months 13-15 Deliverables

☐ **Complete Dataset**

- 1000+ cases with validated CIRF scores
- Geographic representation across 20+ countries
- Balanced outcome distribution maintained
- Data validation and cleaning complete

Months 16-18 Deliverables

☐ **Comprehensive Statistical Analysis**

- All planned statistical tests completed
- Hypothesis testing results finalized
- Predictive models validated and documented
- Effect sizes and practical significance assessed

Phase 5 Milestones (Months 19-24)

Months 19-21 Deliverables

☐ **Research Synthesis**

- Findings interpreted in theoretical context
- Policy and practice implications identified
- Academic publication drafts prepared
- Presentation materials developed

Months 22-24 Deliverables

☐ **Dissertation Completion**

- Complete dissertation draft

- Committee feedback incorporated
- Defense presentation prepared
- Publication strategy finalized

Critical Success Metrics by Phase

Phase 1 Success Criteria

- **Data Quality:** $\geq 80\%$ of cases have high quality scores
- **Methodology Reliability:** Inter-rater reliability ≥ 0.80
- **Technical Performance:** Automated systems operational
- **Statistical Foundation:** Preliminary results support hypotheses

Phase 2 Success Criteria

- **Sample Size:** 200 validated cases collected
- **NLP Accuracy:** $\geq 75\%$ agreement with manual scoring
- **Geographic Coverage:** Cases from ≥ 10 countries
- **Statistical Power:** Sufficient for preliminary hypothesis testing

Phase 3 Success Criteria

- **Scale Achievement:** 600+ cases with quality maintained
- **System Reliability:** $< 5\%$ data collection failure rate
- **Model Performance:** Predictive accuracy $\geq 70\%$
- **Research Progress:** Methodology chapter drafted

Phase 4 Success Criteria

- **Complete Dataset:** 1000+ cases meeting quality standards
- **Statistical Completion:** All planned analyses executed
- **Effect Detection:** Significant relationships identified
- **Practical Significance:** Results have real-world implications

Phase 5 Success Criteria

- **Dissertation Defense:** Successfully defend research
- **Academic Contribution:** At least 2 publications submitted

- **Practical Impact:** Framework validated for practitioners
- **Future Research:** Clear research agenda established

Risk Mitigation Strategies

Technical Risks

- **Data Collection Failure:** Multiple source backups, manual collection protocols
- **NLP Performance:** Human validation samples, confidence thresholds
- **Database Issues:** Regular backups, cloud redundancy

Research Risks

- **Sample Size Shortfall:** Continuous monitoring, alternative sources identified
- **Statistical Power:** Power analyses at each phase, sample size adjustments
- **Hypothesis Support:** Multiple analytical approaches, exploratory analyses

Timeline Risks

- **Scope Creep:** Strict milestone adherence, regular progress reviews
- **Technical Delays:** Buffer time built in, parallel development tracks
- **External Dependencies:** Alternative approaches prepared

Quality Assurance Framework

Continuous Monitoring

- **Weekly Progress Reviews:** Track against milestones
- **Monthly Data Quality Audits:** Validate collection and scoring
- **Quarterly Statistical Reviews:** Assess analytical progress
- **Semi-Annual Comprehensive Reviews:** Full project assessment

Validation Checkpoints

- **Phase 1:** Methodology validation before scaling
- **Phase 2:** Mid-scale validation before production
- **Phase 3:** Production validation before final push
- **Phase 4:** Final validation before synthesis
- **Phase 5:** Comprehensive validation before defense

