

# InfraRader AI

Data Governance & Sourcing Plan

## Data Governance & Sourcing Plan

Comprehensive Framework for Data Management and Acquisition

**Confidential & Proprietary**

Date: October 21, 2025

Contents

<b>1</b>	<b>Executive Summary</b>	<b>2</b>
1.1	Key Objectives . . . . .	2
<b>2</b>	<b>Data Governance Framework</b>	<b>2</b>
2.1	Governance Principles . . . . .	2
2.2	Data Quality Standards . . . . .	2
<b>3</b>	<b>Data Sourcing Strategy</b>	<b>2</b>
3.1	Primary Data Sources . . . . .	2
3.2	Secondary Data Sources . . . . .	3
<b>4</b>	<b>Data Acquisition Workflow</b>	<b>3</b>
4.1	Discovery Phase . . . . .	3
4.2	Collection Phase . . . . .	3
4.3	Processing Phase . . . . .	4
<b>5</b>	<b>Data Privacy and Compliance</b>	<b>4</b>
5.1	Regulatory Framework . . . . .	4
5.2	Data Security Measures . . . . .	4
<b>6</b>	<b>Strategic Data Partnerships</b>	<b>4</b>
6.1	Government Partnerships . . . . .	4
6.2	Industry Partnerships . . . . .	4
6.3	Academic Partnerships . . . . .	4
<b>7</b>	<b>Implementation Timeline</b>	<b>5</b>
7.1	Phase 1: Foundation (Months 1-3) . . . . .	5
7.2	Phase 2: Expansion (Months 4-6) . . . . .	5
7.3	Phase 3: Optimization (Months 7-12) . . . . .	5
<b>8</b>	<b>Success Metrics</b>	<b>5</b>
8.1	Data Quality Metrics . . . . .	5
8.2	Operational Metrics . . . . .	5
<b>9</b>	<b>Risk Management</b>	<b>6</b>
9.1	Data Quality Risks . . . . .	6
9.2	Legal and Compliance Risks . . . . .	6
9.3	Technical Risks . . . . .	6
<b>10</b>	<b>Conclusion</b>	<b>6</b>

## 1 Executive Summary

---

This document outlines InfraRader AI's comprehensive data governance and sourcing strategy, designed to ensure high-quality, reliable, and legally compliant data acquisition for our AI-powered infrastructure intelligence platform.

### 1.1 Key Objectives

- Establish robust data quality standards and validation processes
- Implement comprehensive data sourcing strategies across multiple channels
- Ensure regulatory compliance and data privacy protection
- Create scalable data acquisition and processing workflows
- Build strategic partnerships for enhanced data access

## 2 Data Governance Framework

---

### 2.1 Governance Principles

1. **Data Quality:** All data must meet minimum quality thresholds before processing
2. **Transparency:** Clear documentation of data sources and processing methods
3. **Security:** Robust protection of sensitive and proprietary information
4. **Compliance:** Adherence to all applicable regulations and standards
5. **Scalability:** Framework must support rapid growth and expansion

### 2.2 Data Quality Standards

- **Accuracy:** Minimum 95% accuracy for critical project data
- **Completeness:** All required fields populated for active projects
- **Timeliness:** Data updates within 24-48 hours of availability
- **Consistency:** Standardized formats and terminology across sources
- **Validity:** Data conforms to predefined schemas and constraints

## 3 Data Sourcing Strategy

---

### 3.1 Primary Data Sources

#### 1. Government Databases

- Ministry of Infrastructure websites
- Public procurement portals
- Regulatory filings and permits
- Official project announcements

## 2. Industry Publications

- Construction industry magazines
- Infrastructure trade publications
- Professional association reports
- Industry conference proceedings

## 3. Financial Institutions

- Development bank project databases
- Investment fund announcements
- Loan and financing disclosures
- Public-private partnership records

## 4. Satellite Imagery

- Commercial satellite providers
- Government earth observation data
- Aerial photography services
- Drone imagery partnerships

### 3.2 Secondary Data Sources

- **Social Media:** Company announcements, project updates
- **News Media:** Project coverage, industry analysis
- **Academic Research:** Infrastructure studies, market analysis
- **Consulting Reports:** Industry insights, market intelligence

## 4 Data Acquisition Workflow

---

### 4.1 Discovery Phase

1. **Source Identification:** Automated discovery of new data sources
2. **Source Evaluation:** Assessment of data quality and reliability
3. **Access Negotiation:** Legal agreements and data licensing
4. **Integration Planning:** Technical integration requirements

### 4.2 Collection Phase

1. **Automated Harvesting:** Scheduled data collection from APIs and web scraping
2. **Manual Curation:** Human verification and enhancement
3. **Quality Validation:** Automated and manual quality checks
4. **Data Enrichment:** Additional context and metadata addition

### 4.3 Processing Phase

1. **Data Cleaning:** Standardization and normalization
2. **AI Analysis:** LLM and CV processing for insights
3. **Cross-Validation:** Multi-source verification
4. **Confidence Scoring:** Reliability assessment

## 5 Data Privacy and Compliance

---

### 5.1 Regulatory Framework

- **GDPR Compliance:** European data protection standards
- **CCPA Compliance:** California consumer privacy requirements
- **Local Regulations:** Country-specific data protection laws
- **Industry Standards:** Infrastructure sector compliance requirements

### 5.2 Data Security Measures

- **Encryption:** End-to-end encryption for data transmission and storage
- **Access Control:** Role-based permissions and authentication
- **Audit Trails:** Comprehensive logging of data access and modifications
- **Data Retention:** Automated data lifecycle management

## 6 Strategic Data Partnerships

---

### 6.1 Government Partnerships

- **Ministry Agreements:** Direct access to official project data
- **Regulatory Bodies:** Authorized access to permits and approvals
- **Public Agencies:** Infrastructure development authorities

### 6.2 Industry Partnerships

- **Data Providers:** Commercial data aggregators and providers
- **Technology Partners:** Satellite imagery and geospatial data companies
- **Industry Associations:** Professional and trade organizations

### 6.3 Academic Partnerships

- **Research Institutions:** Access to academic research and studies
- **Think Tanks:** Policy and market analysis organizations
- **International Organizations:** Development and infrastructure agencies

## 7 Implementation Timeline

---

### 7.1 Phase 1: Foundation (Months 1-3)

- Establish data governance framework
- Implement basic data quality standards
- Set up initial data sources
- Develop core collection workflows

### 7.2 Phase 2: Expansion (Months 4-6)

- Add secondary data sources
- Implement advanced quality validation
- Establish strategic partnerships
- Deploy AI-powered processing

### 7.3 Phase 3: Optimization (Months 7-12)

- Refine data quality metrics
- Optimize collection workflows
- Expand partnership network
- Implement advanced analytics

## 8 Success Metrics

---

### 8.1 Data Quality Metrics

- **Accuracy Rate:** Percentage of verified accurate data points
- **Completeness Score:** Percentage of required fields populated
- **Freshness Index:** Average age of data updates
- **Source Diversity:** Number of independent data sources per project

### 8.2 Operational Metrics

- **Collection Volume:** Amount of data collected per time period
- **Processing Speed:** Time from collection to analysis completion
- **Cost Efficiency:** Cost per data point acquired and processed
- **Partnership Growth:** Number of active data partnerships

## 9 Risk Management

---

### 9.1 Data Quality Risks

- **Risk:** Inaccurate or outdated information
- **Mitigation:** Multi-source validation and confidence scoring

### 9.2 Legal and Compliance Risks

- **Risk:** Data privacy violations or regulatory non-compliance
- **Mitigation:** Comprehensive legal review and compliance framework

### 9.3 Technical Risks

- **Risk:** Data source unavailability or format changes
- **Mitigation:** Diversified sourcing and robust error handling

## 10 Conclusion

---

This data governance and sourcing plan provides a comprehensive framework for ensuring high-quality, reliable, and compliant data acquisition for InfraRader AI. Through systematic implementation of these strategies, we will build a robust data foundation that supports our AI-powered infrastructure intelligence platform and enables sustainable growth and market leadership.