

Environmental Impact Investing Monitor (EIIM) - Product Requirements Document

Executive Summary

Product Name: Environmental Impact Investing Monitor (EIIM)

Version: 1.0

Target Users: Environmental impact investors, consultants, and analysts

Development Approach: Free and open-source tools only

Core Functionality: Daily news briefings, carbon price tracking with historical context, metrics dashboard, and archive system

1. Product Overview

1.1 Problem Statement

Environmental impact investors need real-time intelligence on market developments, carbon pricing trends, and ecosystem health metrics. Current solutions are fragmented, expensive, and lack comprehensive coverage of emerging markets like biodiversity credits.

1.2 Solution

A comprehensive web application that aggregates news from 47+ sources, tracks multiple carbon pricing streams with historical analysis, monitors 18 key ecosystem metrics, and provides AI-powered insights through OpenRouter API integration.

1.3 Success Metrics

- Daily active users engaging with morning briefs (target: 80%+ open rate within 2 hours)
- Historical data retention covering 2+ years
- Sub-3 second page load times
- 99% uptime for data collection services
- User engagement with archived content

2. Technical Architecture

2.1 Technology Stack (Free Tools Only)

Frontend:

- React 18+ with TypeScript

- Tailwind CSS for styling
- Chart.js or Recharts for data visualization
- React Router for navigation
- React Query for state management

Backend:

- Node.js with Express
- PostgreSQL database (upgradeable from SQLite)
- Node-cron for scheduled tasks
- Cheerio for web scraping
- Axios for API calls

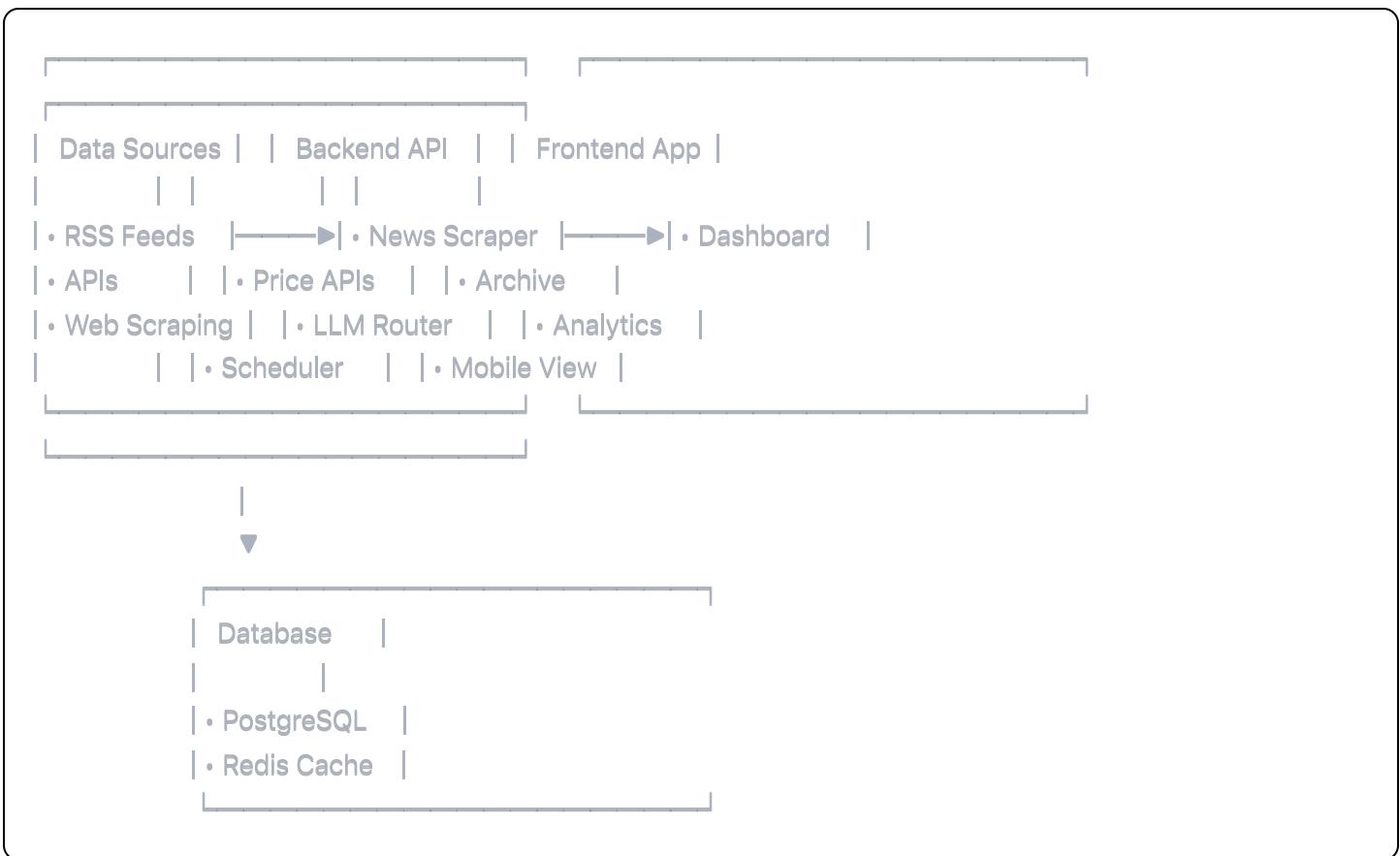
Deployment:

- Frontend: Vercel (free tier)
- Backend: Railway (free tier) or Render (free tier)
- Database: Railway PostgreSQL (free tier)
- File Storage: GitHub repository for static assets

Development Tools:

- GitHub for version control
- GitHub Actions for CI/CD
- ESLint + Prettier for code quality
- Jest for testing

2.2 System Architecture



2.3 Data Sources Architecture

News Aggregation:

- RSS feeds from identified sources
- Web scraping for sources without RSS
- Scheduled collection every 4 hours
- Duplicate detection and content normalization

Carbon Price APIs (Free/Public):

- World Bank Carbon Pricing Dashboard API
- Alpha Vantage (500 requests/day free)
- Yahoo Finance API (commodity data)
- Quandl (500 calls/day free)
- Custom scrapers for exchange websites

Historical Data Storage:

- Daily price snapshots
- Weekly aggregations

- Monthly trend analysis
- 5-year retention minimum

3. Core Features

3.1 Morning News Brief System

Data Collection Pipeline:

RSS/Web Scraping → Content Processing → AI Summarization → Categorization → Storage

Implementation Details:

- Scraping Frequency: Every 4 hours (6 collections/day)
- Content Processing: Extract title, content, date, source, URL
- Deduplication: Content hashing to prevent duplicates
- Categorization: ML classification into predefined categories
- AI Integration: OpenRouter API for content summarization

Categories:

1. Venture Capital & Early Stage
2. Public Markets & Green Bonds
3. Carbon Markets & Pricing
4. Policy & Regulation
5. Technology Breakthroughs
6. Biodiversity & Nature Credits
7. ESG & Sustainability

Brief Generation Logic:

- Priority scoring based on source credibility, recency, and category
- AI-powered summarization of top 15-20 articles
- Regional focus weighting (60% US/North America, 40% global)
- Investment stage relevance (70% early-stage, 30% public markets)

3.2 Carbon Price Tracking Dashboard

Multi-Price Architecture:

typescript

```
const carbonPrices = {
  compliance: {
    eu_ets: { current: 0, change: 0, volume: 0 },
    california: { current: 0, change: 0, volume: 0 },
    rrgi: { current: 0, change: 0, volume: 0 },
    uk_ets: { current: 0, change: 0, volume: 0 }
  },
  voluntary: {
    verra_vcs: { current: 0, change: 0, volume: 0 },
    gold_standard: { current: 0, change: 0, volume: 0 },
    art_trees: { current: 0, change: 0, volume: 0 },
    puro_removal: { current: 0, change: 0, volume: 0 }
  },
  indices: {
    ice_global: { current: 0, change: 0 },
    trove_research: { current: 0, change: 0 }
  }
};
```

Historical Context Features:

- Trend Analysis: 30-day, 90-day, 1-year percentage changes
- Volatility Metrics: Standard deviation, VaR calculations
- Correlation Matrix: Cross-market relationships
- Event Impact: Policy announcements, market shocks
- Seasonal Patterns: Monthly/quarterly cyclical trends

Price Collection Logic:

typescript

```

// Scheduled every 15 minutes during market hours
const collectPrices = async () => {
  const sources = [
    { name: 'worldbank', endpoint: '/carbon-pricing-dashboard/api' },
    { name: 'alpha_vantage', endpoint: '/query?function=COMMODITY' },
    { name: 'yahoo_finance', endpoint: '/v8/finance/chart/CO2-EUR.EX' }
  ];

  for (const source of sources) {
    try {
      const data = await fetchPriceData(source);
      await storePriceData(data);
      await calculateHistoricalMetrics(data);
    } catch (error) {
      console.error(`Failed to collect from ${source.name}:`, error);
    }
  }
};

```

3.3 Key Metrics Dashboard

18 Core Metrics Implementation:

1. Climate Tech Venture Funding Volume (quarterly)

- Data: PwC State of Climate Tech, HolonIQ
- API: Custom scraper + manual updates
- Historical: 5-year trend analysis

2. Climate Tech Deal Count (monthly)

- Data: CTVC database, Pitchbook (free tier)
- Tracking: Series A-C focusing on North America

3. Average Deal Size by Stage (quarterly)

- Calculation: Volume/Count by investment stage
- Segmentation: Pre-seed, Seed, Series A, B, C+

4. New Climate Fund Formation (quarterly)

- Data: Preqin (free data), press releases
- Tracking: AUM, focus areas, geographic concentration

5. Green Bond Issuance Volume (monthly)

- Data: Climate Bonds Initiative (free reports)

- Categories: Corporate, municipal, sovereign, supranational

6. ESG Fund Flows (monthly)

- Data: Morningstar Sustainable Investing (free reports)
- Regional breakdown: US, Europe, Asia-Pacific

7. Clean Energy Stock Index Performance (daily)

- Data: Yahoo Finance API (free)
- Indices: ICLN, PBW, QCLN vs S&P 500

8. Carbon Credit Market Volume (monthly)

- Data: Ecosystem Marketplace, AlliedOffsets
- Split: Compliance vs voluntary markets

9. Climate Patent Filings (quarterly)

- Data: WIPO Green Inventory (free access)
- Categories: Clean energy, carbon capture, efficiency

10. Corporate Net-Zero Commitments (monthly)

- Data: Science Based Targets initiative (free)
- Tracking: New commitments, target dates, progress

11. Renewable Energy Capacity Additions (quarterly)

- Data: IRENA Global Energy Transformation (free)
- Breakdown: Solar, wind, storage, other

12. Carbon Removal Deployment (quarterly)

- Data: State of CDR report (free)
- Methods: DAC, BECCS, enhanced weathering, forestry

13. Environmental Policy Stringency Index (annual)

- Data: OECD (free access)
- Countries: G20 + key emerging markets

14. Government Green Investment as % GDP (annual)

- Data: IEA Energy Investment report (free summary)
- Regional comparison and trend analysis

15. Biodiversity Credit Market Volume (quarterly)

- Data: Ecosystem Marketplace, Pollination Group
- Tracking: Nascent market development

16. Blue Carbon Project Pipeline (quarterly)

- Data: Blue Carbon Initiative, conservation orgs
- Project count, hectares, credit potential

17. Water Credit Market Activity (quarterly)

- Data: EPA water quality trading database
- Regional programs: Ohio River, Chesapeake Bay

18. Plastic Credit Market Volume (quarterly)

- Data: Verra Plastic Program, PCX Solutions
- Tracking: Credit issuance and retirement

3.4 Archive System Architecture

Database Schema:

sql

```
-- News Articles Archive
CREATE TABLE news_articles (
    id SERIAL PRIMARY KEY,
    title VARCHAR(500) NOT NULL,
    content TEXT,
    summary TEXT, -- AI-generated
    source VARCHAR(100),
    url VARCHAR(500),
    published_date TIMESTAMP,
    scraped_date TIMESTAMP DEFAULT NOW(),
    category VARCHAR(50),
    tags TEXT[], -- Array of tags
    priority_score INTEGER,
    sentiment_score DECIMAL(3,2) -- -1 to 1
);
```

```
-- Daily Briefs Archive
CREATE TABLE daily_briefs (
    id SERIAL PRIMARY KEY,
    brief_date DATE UNIQUE,
    content TEXT, -- Full brief HTML/Markdown
    article_count INTEGER,
    top_categories TEXT[],
    generated_at TIMESTAMP DEFAULT NOW(),
    ai_model_used VARCHAR(50) -- Track which model generated
);
```

```
-- Carbon Prices Historical
CREATE TABLE carbon_prices (
    id SERIAL PRIMARY KEY,
    market VARCHAR(50), -- eu_ets, california, etc.
    price DECIMAL(10,2),
    volume BIGINT,
    currency VARCHAR(3),
    timestamp TIMESTAMP,
    data_source VARCHAR(50)
);
```

```
-- Metrics Historical
CREATE TABLE ecosystem_metrics (
    id SERIAL PRIMARY KEY,
    metric_name VARCHAR(100),
    value DECIMAL(15,2),
```

```
unit VARCHAR(50),
period_start DATE,
period_end DATE,
geography VARCHAR(50),
data_source VARCHAR(100),
recorded_at TIMESTAMP DEFAULT NOW()
);
```

Archive Features:

- Full-text search across all historical briefs
- Date range filtering with calendar interface
- Category filtering and tag-based search
- Export functionality (PDF, CSV, JSON)
- Comparative analysis between time periods
- Bookmark system for important briefs/articles

3.5 AI Integration via OpenRouter

Configuration System:

```
typescript

const openRouterConfig = {
  baseURL: 'https://openrouter.ai/api/v1',
  defaultModel: 'anthropic/clause-3-haiku', // Cost-effective default
  models: {
    summarization: 'anthropic/clause-3-haiku',
    analysis: 'anthropic/clause-3-sonnet',
    research: 'anthropic/clause-3-opus', // For complex analysis
    coding: 'deepseek/deepseek-coder',
    alternative: 'meta-llama/llama-3.1-8b-instruct' // Backup/cost option
  },
  rateLimits: {
    requestsPerMinute: 20,
    tokensPerMinute: 100000
  }
};
```

AI Use Cases:

1. Content Summarization: Reduce 500-word articles to 100-word summaries

2. Brief Generation: Synthesize top stories into coherent morning brief
3. Trend Analysis: Identify patterns in historical data
4. Sentiment Analysis: Gauge market sentiment from news content
5. Anomaly Detection: Flag unusual price movements or metric changes
6. Query Processing: Natural language search of archive
7. Report Generation: Automated weekly/monthly reports

Prompt Templates:

```
typescript

const prompts = {
  summarizeArticle: `Summarize this environmental finance article in exactly 100 words, focusing on investment in renewable energy projects. ${article}`,

  generateBrief: `Create a daily morning brief from these {count} environmental impact investing articles. Structure it as a bullet-point summary of key trends and opportunities. ${articles}`,

  analyzeTrend: `Analyze this carbon pricing data for trends, patterns, and investment implications. Include volatility and projected impacts on different sectors. ${data}`,

  detectAnomalies: `Review these metrics for anomalies or significant changes that warrant investor attention:\n\n${metrics}`,
};
```

4. Data Collection Strategy

4.1 News Sources Implementation

Tier 1 - Daily Priority Sources (7):

1. Environmental Finance (environmental-finance.com)
2. Carbon Pulse (carbon-pulse.com)
3. Bloomberg Green (bloomberg.com/green)
4. ESG Today (esgtoday.com)
5. Climate Tech VC (climatetechvc.substack.com)
6. GreenBiz (greenbiz.com)
7. Sustainable Finance (sustainable-finance.hsbc.com)

Tier 2 - Weekly Deep Dive Sources (15): 8. Climate Bonds Initiative (climatebonds.net) 9. Ecosystem Marketplace (ecosystemmarketplace.com) 10. Climate Policy Initiative (climatepolicyinitiative.org) 11. Rocky Mountain Institute (rmi.org) 12. World Bank Climate Change (worldbank.org/climate) 13. UNEP Finance Initiative (unepfi.org) 14. Ceres (ceres.org) 15. CDP (cdp.net) 16. Climate Action 100+ (climateaction100.org)

(climateaction100.org) 17. Task Force on Climate-related Financial Disclosures (fsb-tcfd.org) 18. Network for Greening the Financial System (bis.org/ngfs) 19. International Energy Agency (iea.org) 20. International Renewable Energy Agency (irena.org) 21. Global Sustainable Investment Alliance (gsi-alliance.org) 22. Principles for Responsible Investment (unpri.org)

Tier 3 - Monthly Strategic Sources (25): 23. McKinsey Sustainability (mckinsey.com/sustainability)
24. Boston Consulting Group Climate (bcg.com/climate) 25. Deloitte Sustainability
(deloitte.com/sustainability) 26. PwC Sustainability (pwc.com/sustainability) 27. KPMG ESG
(kpmg.com/esg) 28. Morningstar Sustainable Investing (morningstar.com/sustainable-investing) 29. S&P Global Sustainable1 (spglobal.com/sustainable1) 30. MSCI ESG Research (msci.com/esg) 31. Refinitiv ESG (refinitiv.com/esg) 32. RepRisk (reprisk.com) 33. Sustainalytics (sustainalytics.com) 34. Climate Trace (climatetrace.org) 35. Climate Central (climatecentral.org) 36. Carbon Brief
(carbonbrief.org) 37. Inside Climate News (insideclimanews.org) 38. Climate Home News
(climatechangenews.com) 39. Mongabay (mongabay.com) 40. Yale Environment 360 (e360.yale.edu)
41. The Guardian Environment (theguardian.com/environment)
42. Financial Times Climate Capital (ft.com/climate-capital) 43. Reuters Sustainability
(reuters.com/sustainability) 44. Wall Street Journal ESG (wsj.com/esg) 45. Axios Climate
(axios.com/climate) 46. Politico Energy (politico.com/energy) 47. Climate Wire (climatewire.com)

RSS-Based Sources (Priority 1):

```
typescript

const rssSources = [
  'https://www.environmental-finance.com/rss',
  'https://carbon-pulse.com/feed/',
  'https://www.bloomberg.com/green/feed',
  'https://www.esgtoday.com/feed/',
  'https://climatetechvc.substack.com/feed'
  // ... additional RSS feeds
];
```

Web Scraping Sources (Priority 2):

```
typescript
```

```
const scrapingSources = [
  {
    name: 'Climate Bonds Initiative',
    url: 'https://www.climatebonds.net/market/data/',
    selector: '.news-item',
    frequency: 'daily'
  },
  {
    name: 'CTVC Database',
    url: 'https://www.ctvc.co/deals',
    selector: 'deal-item',
    frequency: 'weekly'
  }
  // ... additional scraping targets
];
```

Error Handling Strategy:

- Retry failed requests with exponential backoff
- Fallback to cached content if source unavailable
- Alert system for sustained source failures
- Manual backup procedures for critical sources

4.2 API Rate Limiting Strategy

OpenRouter Management:

typescript

```

class OpenRouterManager {
  constructor() {
    this.requestQueue = [];
    this.isProcessing = false;
    this.rateLimiter = new RateLimiter(20, 60000); // 20 req/min
  }

  async makeRequest(prompt, model = 'claude-3-haiku') {
    return this.rateLimiter.schedule(async () => {
      const response = await fetch('https://openrouter.ai/api/v1/chat/completions', {
        method: 'POST',
        headers: {
          'Authorization': `Bearer ${process.env.OPENROUTER_API_KEY}`,
          'Content-Type': 'application/json'
        },
        body: JSON.stringify({
          model: `anthropic/${model}`,
          messages: [{ role: 'user', content: prompt }],
          max_tokens: 1000
        })
      });
      return response.json();
    });
  }
}

```

5. Development Phases

5.1 Phase 1: MVP Core (Weeks 1-4)

Sprint 1-2: Foundation

- Set up development environment and CI/CD
- Implement basic React frontend with routing
- Create SQLite database with core schemas
- Build news scraping pipeline for top 10 sources
- Implement basic archive storage

Sprint 3-4: Core Features

- Develop daily brief generation system
- Integrate OpenRouter API for summarization

- Create carbon price collection for 5 major markets
- Build historical price storage and basic charts
- Implement simple archive viewing

5.2 Phase 2: Enhanced Features (Weeks 5-8)

Sprint 5-6: Advanced Dashboard

- Expand to all 47 news sources
- Implement comprehensive carbon price tracking
- Add historical context and trend analysis
- Build metrics dashboard for top 10 KPIs
- Create search functionality for archives

Sprint 7-8: Intelligence Layer

- Advanced AI integration for brief generation
- Sentiment analysis and trend detection
- Anomaly detection for price movements
- Export functionality and reporting
- Performance optimization

5.3 Phase 3: Production Ready (Weeks 9-12)

Sprint 9-10: Scaling & Reliability

- Migrate to PostgreSQL on Railway
- Implement comprehensive error handling
- Add monitoring and alerting systems
- Optimize for production deployment
- Complete all 18 metrics integration

Sprint 11-12: User Experience

- Polish UI/UX with professional design
- Advanced filtering and search capabilities
- Mobile responsiveness
- User preferences and customization

- Documentation and onboarding

6. Security & Privacy

6.1 API Key Management

- Environment variables for all sensitive data
- Separate keys for development/production
- Key rotation procedures documented
- No hardcoded credentials in repository

6.2 Data Privacy

- No personal user data collection initially
- GDPR-compliant data handling procedures
- Clear data retention policies
- Secure data transmission (HTTPS only)

6.3 Rate Limiting & Abuse Prevention

- Request throttling on public endpoints
- IP-based rate limiting for scraping
- Graceful degradation under high load
- Monitoring for unusual usage patterns

7. Monitoring & Analytics

7.1 Application Monitoring

```
typescript

const monitoring = {
  uptime: 'UptimeRobot (free tier)',
  errors: 'Sentry (free tier)',
  performance: 'Built-in timing metrics',
  logs: 'Railway built-in logging'
};
```

7.2 Business Metrics

- Daily active brief views

- Archive search queries
- API response times
- Data collection success rates
- User engagement patterns

7.3 Alerting System

- Failed data collection alerts
- API quota warnings
- Performance degradation notices
- Critical error notifications

8. Deployment & Infrastructure

8.1 Free Tier Resource Allocation

Vercel (Frontend):

- 100GB bandwidth/month
- Unlimited static hosting
- Serverless functions for API routes

Railway (Backend & Database):

- \$5/month credit (sufficient for development)
- PostgreSQL database included
- Auto-scaling within limits

GitHub (Storage & CI/CD):

- 2GB repository storage
- 2,000 minutes/month Actions
- Unlimited public repositories

8.2 CI/CD Pipeline

```
yaml
```

```

# .github/workflows/deploy.yml
name: Deploy Application
on:
  push:
    branches: [main]
jobs:
  deploy-frontend:
    runs-on: ubuntu-latest
    steps:
      - uses: actions/checkout@v3
      - uses: actions/setup-node@v3
      - run: npm ci && npm run build
      - uses: amondnet/vercel-action@v20

  deploy-backend:
    runs-on: ubuntu-latest
    steps:
      - uses: actions/checkout@v3
      - uses: actions/setup-node@v3
      - run: npm ci
      - uses: railway-deploy-action@v1

```

8.3 Environment Configuration

```

typescript

// config/environment.js
module.exports = {
  development: {
    database: 'sqlite:./dev.db',
    openrouter_api_key: process.env.OPENROUTER_API_KEY_DEV,
    scraping_interval: 3600000 // 1 hour
  },
  production: {
    database: process.env.DATABASE_URL,
    openrouter_api_key: process.env.OPENROUTER_API_KEY_PROD,
    scraping_interval: 900000 // 15 minutes
  }
};

```

9. Testing Strategy

9.1 Unit Testing

- Jest for JavaScript testing
- React Testing Library for component tests
- Supertest for API endpoint testing
- 80% code coverage minimum

9.2 Integration Testing

- Database operations testing
- API integration testing
- Scraping pipeline validation
- OpenRouter API mocking

9.3 End-to-End Testing

- Cypress for critical user journeys
- Morning brief generation workflow
- Archive search and filtering
- Historical data visualization

10. Documentation Requirements

10.1 Technical Documentation

- API documentation with OpenAPI spec
- Database schema documentation
- Deployment procedures
- Troubleshooting guides

10.2 User Documentation

- Getting started guide
- Feature walkthrough
- FAQ section
- Changelog maintenance

11. Success Criteria & KPIs

11.1 Technical KPIs

- 99% uptime for data collection
- <3 second page load times
- <5 minute brief generation time
- 100% API test coverage

11.2 Business KPIs

- Daily brief engagement rate >70%
- Archive search usage growth
- Historical data retention 2+ years
- Zero data loss incidents

11.3 User Experience KPIs

- <2 second time to first meaningful paint
- Mobile responsiveness score >90
- Accessibility compliance (WCAG 2.1)
- User session duration growth

12. Risk Assessment & Mitigation

12.1 Technical Risks

- **Risk:** API rate limits exceeded
Mitigation: Implement queue system, multiple API keys, fallback sources
- **Risk:** Web scraping blocked
Mitigation: Rotate user agents, respect robots.txt, implement delays
- **Risk:** Database storage limits
Mitigation: Automated cleanup, data archiving, compression

12.2 Business Risks

- **Risk:** News source access revoked
Mitigation: Diversified source portfolio, backup sources identified
- **Risk:** OpenRouter API changes
Mitigation: Abstract API layer, multiple model support

- **Risk:** Free tier limitations exceeded

Mitigation: Usage monitoring, graceful degradation, upgrade path

13. Future Enhancements (Post-MVP)

13.1 Advanced Features

- Machine learning for trend prediction
- Custom alert system for price thresholds
- Portfolio integration for impact assessment
- Collaborative features for team usage
- Advanced visualization with D3.js

13.2 Monetization Options

- Premium tier with real-time data
- Custom reporting services
- API access for third parties
- White-label solutions
- Consulting services integration

13.3 Scaling Considerations

- Microservices architecture migration
- Dedicated database instances
- CDN implementation for global performance
- Advanced caching strategies
- Load balancing for high availability

14. API Endpoints Documentation

Base URL: <https://api.eiim.app/v1>

Core Endpoints:

```
GET /briefs           # Get daily briefs with pagination
GET /briefs/:date    # Get specific date brief
GET /carbon-prices   # Current carbon prices
GET /carbon-prices/historical # Historical price data
GET /metrics          # Current ecosystem metrics
GET /metrics/historical # Historical metrics data
GET /search           # Search archived content
POST /alerts          # Create price/metric alerts
```

15. Database Optimization Queries

sql

```
-- Optimize news search performance
CREATE INDEX idx_news_date_category ON news_articles(published_date, category);
CREATE INDEX idx_news_fulltext ON news_articles USING gin(to_tsvector('english', title || ' ' || content));

-- Optimize carbon price queries
CREATE INDEX idx_carbon_prices_market_timestamp ON carbon_prices(market, timestamp);
CREATE INDEX idx_carbon_prices_timestamp_desc ON carbon_prices(timestamp DESC);

-- Optimize metrics queries
CREATE INDEX idx_metrics_name_period ON ecosystem_metrics(metric_name, period_end);
```

16. Deployment Checklist

Pre-deployment:

- Environment variables configured
- Database migrations run
- SSL certificates valid
- Monitoring alerts configured
- Backup procedures tested

Post-deployment:

- Health checks passing
- Data collection pipelines active
- Error tracking functional
- Performance monitoring active
- User acceptance testing complete

This PRD provides comprehensive guidance for building the Environmental Impact Investing Monitor using exclusively free and publicly available tools, with clear technical specifications for automated development implementation.