

PulseNet - Progress Tracker

Current Status & Development Log

Project Name: PulseNet (Real-Time Global Emotion Map)

Start Date: October 18, 2025

Current Sprint: MVP Development (Week 1-2)

Last Updated: October 20, 2025

Team Lead: Victor

OVERALL PROGRESS: 55%



55% Complete



CURRENT PROJECT STRUCTURE



emotion-map-project/

└── backend/	
└── app.py	✓ Main Flask application
└── config.py	✓ Configuration management
└── data_collection/	
└── __init__.py	✓ Package init
└── rss_collector.py	✓ RSS feeds (733 posts)
└── reddit_collector.py	✓ Reddit API (175 posts)
└── news_collector.py	✓ NewsData.io (500+ posts)
└── twitter_collector.py	⌚ NOT CREATED YET
└── processing/	
└── __init__.py	✓ Package init (empty)
└── emotion_analyzer.py	⌚ NOT CREATED YET
└── location_extractor.py	⌚ NOT CREATED YET
└── aggregator.py	⌚ NOT CREATED YET
└── database/	
└── __init__.py	✓ Package init
└── init_db.py	✓ Database initialization
└── db_manager.py	✓ CRUD operations
└── scheduler/	
└── __init__.py	✓ Package init
└── background_tasks.py	⌚ PARTIAL (only RSS)
└── routes/	
└── __init__.py	✓ Package init
└── api_routes.py	✓ 5 API endpoints
└── frontend/	
└── static/	
└── css/	
└── style.css	⌚ NOT CREATED YET
└── js/	
└── map.js	⌚ NOT CREATED YET
└── dashboard.js	⌚ NOT CREATED YET
└── api.js	⌚ NOT CREATED YET
└── images/	📁 Empty folder

	 BASIC PLACEHOLDER
└── templates/	
└── index.html	
└── tests/	
├── test_api.py	 NOT CREATED YET
├── test_collectors.py	 NOT CREATED YET
└── test_processing.py	 NOT CREATED YET
└── .env	 Environment variables
└── .gitignore	 Git ignore rules
└── requirements.txt	 Python dependencies
└── test_db.py	 Database testing script
└── README.md	 NEEDS UPDATE
└── data_collection.log	 Auto-generated log file
└── skipped_subreddits.log	 Auto-generated log file
└── emotion_map.db	 SQLite database (908+ posts)

Legend:

-  Complete and working
-  Partially complete
-  Not started
-  Empty folder

COMPLETED TASKS

Infrastructure (100% Complete) - Week 1

1. Development Environment

- Installed Python 3.14.0
- Installed UV package manager
- Created virtual environment
- Installed all dependencies
- Set up Git version control
- Created GitHub repository

Repository: <https://github.com/vixtor-e86/emotion-map-project>

2. Project Structure

- Created backend folder structure
- Created frontend folder structure

- Created tests folder
- Set up .gitignore
- Created .env configuration file

3. Database Setup

- Designed database schema
- Created raw_posts table
- Created aggregated_sentiment table (needs emotion columns update)
- Added database indexes
- Implemented DatabaseManager class
- Tested database operations

Database File: emotion_map.db (SQLite)

4. Configuration System

- Created backend/config.py
- Set up environment variables
- Added API key management
- Configured Flask settings

API Keys Configured:

- Reddit API (client ID + secret)
- NewsData.io API key
- Twitter API (pending)

5. Flask Application

- Created main Flask app
- Set up CORS
- Implemented application factory pattern
- Created blueprint structure
- Basic HTML template

6. API Endpoints

- /api/health - Health check
- /api/map-data/<zoom_level> - Map data
- /api/location/<location_name> - Location details
- /api/stats - Global statistics
- /api/trends - Trends endpoint (placeholder)

Status: All endpoints tested and working

Data Collection (80% Complete) - Week 1-2

1. RSS Feed Collector

File: backend/data_collection/rss_collector.py

Status:  WORKING

- Integrated 20 international RSS feeds
- Implemented feed parsing with feedparser
- Added error handling for invalid feeds
- Connected to database
- Country mapping

Performance:

- Sources: 20 RSS feeds
- Success Rate: 75% (15/20 working)
- Data Collected: 733 posts per run
- Coverage: Global (UK, USA, Europe, Asia, Africa, Australia)

2. Reddit Collector

File: backend/data_collection/reddit_collector.py

Status:  WORKING

- Connected to Reddit API (PRAW)
- Implemented 54 global subreddits
- Added rate limit protection
- Connected to database
- Country mapping for subreddits

Performance:

- Subreddits: 54 active communities
- Data Collected: 175 posts per run (5 per subreddit for testing)
- Coverage: 6 continents, 40+ countries
- Rate Limit: Respected (2-4 second delays)

3. News API Collector

File: backend/data_collection/news_collector.py

Status:  WORKING (Just Fixed!)

- Connected to NewsData.io API
- Implemented 32 country coverage
- Added rate limit handling
- Connected to database
- Direct country mapping

Performance:

- Countries: 32 covered
- Expected: 500-800 articles per run
- Rate Limit: 100 requests/day

- Delay: 2 seconds per request

4. Twitter Collector

File: backend/data_collection/twitter_collector.py

Status:  PENDING

- Get Twitter API access
- Implement tweet collection
- Add location extraction
- Connect to database
- Test and validate

Note: Twitter API requires application approval. Team working on this.

IN PROGRESS TASKS

Processing Layer (20% Complete) - Week 2

1. Emotion Analyzer

File: backend/processing/emotion_analyzer.py

Status:  NEEDS CREATION

Recommended Approach:



python

```
# Option 1: VADER + Keyword Mapping (Simplest)  RECOMMENDED
# 1. Use VADER for sentiment baseline
# 2. Map keywords to 5 emotions
# 3. Combine scores for final classification
```

```
EMOTION_KEYWORDS = {
    'joy': ['happy', 'love', 'amazing', 'wonderful', 'excited', 'great'],
    'anger': ['hate', 'angry', 'furious', 'terrible', 'worst', 'outrage'],
    'fear': ['scared', 'afraid', 'worry', 'panic', 'anxiety', 'terrified'],
    'hope': ['hope', 'optimistic', 'believe', 'faith', 'confident', 'better'],
    'calmness': ['calm', 'peace', 'okay', 'fine', 'normal', 'steady']
}
```

Tasks Remaining:

- Create emotion_analyzer.py
- Implement keyword-based classifier
- Integrate VADER sentiment scores
- Add confidence scoring
- Test on sample data
- Update raw_posts with emotion data

Why This Approach:

- Simple to implement (1-2 days)
- No complex ML models needed
- Fast processing (100+ posts/second)
- Easy to debug and adjust
- Good accuracy for MVP (70%+)

2. Location Extractor

File: backend/processing/location_extractor.py

Status:  NEEDS CREATION

Recommended Approach:



python

```
# Three-layer strategy:
# Layer 1: Source-based mapping (Reddit, News API have countries)
# Layer 2: Regex patterns for common mentions
# Layer 3: Manual city/country dictionary (200+ locations)
```

```
LOCATION_PATTERNS = [
    r'in ([A-Z][a-z]+)',
    r'from ([A-Z][a-z]+)',
    r'#([A-Z][a-z]+)',
]
```

```
CITY_COUNTRY_MAP = {
    'Paris': 'France',
    'Lagos': 'Nigeria',
    'Tokyo': 'Japan',
    # ... 200+ cities
}
```

Tasks Remaining:

- Create location_extractor.py
- Build regex patterns
- Create 200+ city/country dictionary
- Add continent mapping
- Integrate geopy as fallback
- Test accuracy
- Update raw_posts with locations

3. Data Aggregator

File: backend/processing/aggregator.py

Status:  NEEDS CREATION

Purpose: Group posts by location and calculate emotion statistics

Tasks Remaining:

- Create aggregator.py
- Implement country-level aggregation
- Implement continent-level aggregation
- Calculate emotion percentages
- Determine dominant emotion
- Save to aggregated_emotions table
- Add time-based filtering

Logic:



python

```
# For each country:
# 1. Count posts per emotion
# 2. Calculate percentages
# 3. Find dominant emotion
# 4. Compute average confidence scores
# 5. Store in aggregated_emotions table
```

4. Background Scheduler

File: backend/scheduler/background_tasks.py

Status:  PARTIALLY COMPLETE

Current State:

- APScheduler configured
- RSS collector integrated

- Reddit collector integration
- News collector integration
- Twitter collector integration
- Processing pipeline
- Error handling
- Logging

Complete Pipeline Needed:



python

```
def complete_data_pipeline():
    # 1. Collect from all sources
    rss_count = collect_rss_data()
    reddit_count = collect_reddit_data()
    news_count = collect_news_data()
    # twitter_count = collect_twitter_data()

    # 2. Process all new posts
    analyze_emotions()
    extract_locations()

    # 3. Aggregate by location
    aggregate_by_country()
    aggregate_by_continent()

    # 4. Cleanup old data
    cleanup_old_data(days=7)
```

PENDING TASKS

Frontend Development (0% Complete) - Week 2-3

1. HTML Structure

File: frontend/templates/index.html

Status: BASIC PLACEHOLDER ONLY

Tasks:

- Create complete HTML structure
- Add global emotion meter section

- Add map container div
- Add dashboard cards section
- Add time slider/replay controls
- Add search bar
- Add filter controls
- Add modal for location details
- Make mobile responsive

2. CSS Styling

File: frontend/static/css/style.css

Status:  NOT STARTED

Tasks:

- Design color scheme (5 emotions)
- Style header and navigation
- Style global emotion meter
- Style map container
- Style dashboard cards
- Style modal/popups
- Add animations
- Add hover effects
- Mobile responsive design
- Dark mode support (optional)

Color Palette:



```
:root {  
  --joy: #FFC107;  
  --anger: #F44336;  
  --fear: #9C27B0;  
  --hope: #4CAF50;  
  --calmness: #2196F3;  
}
```

3. Map Visualization

File: frontend/static/js/map.js

Status:  NOT STARTED

Tasks:

- Initialize Plotly.js choropleth map
- Fetch data from /api/map-data
- Color countries by dominant emotion
- Add hover tooltips
- Add click handlers
- Implement zoom levels (continent/country/city)
- Auto-refresh every 60 minutes
- Add loading states
- Handle errors gracefully

Key Features:

- Interactive map
- Color-coded by emotion
- Click for details
- Smooth transitions

4. Dashboard Components

File: frontend/static/js/dashboard.js

Status:  NOT STARTED

Tasks:

- Create global emotion meter
- Display total posts count
- Show countries tracked
- Display dominant emotion
- Create emotion trend chart
- Add sample posts display
- Implement time filters
- Auto-update statistics

5. API Integration

File: frontend/static/js/api.js

Status:  NOT STARTED

Tasks:

- Create fetch functions for all endpoints
- Add error handling
- Add loading indicators
- Implement caching (optional)
- Add retry logic

Functions Needed:



javascript

- [fetchMapData\(zoomLevel\)](#)
 - [fetchLocationDetails\(location\)](#)
 - [fetchGlobalStats\(\)](#)
 - [fetchTrends\(hours\)](#)
 - [searchEmotions\(query\)](#)
-

Testing (0% Complete)

Unit Tests

Status:  NOT STARTED

Tasks:

- Test database operations
- Test API endpoints
- Test data collectors
- Test emotion analyzer
- Test location extractor
- Test aggregator

Integration Tests

Status:  NOT STARTED

Tasks:

- Test complete data pipeline
- Test frontend-backend integration
- Test error handling
- Test rate limiting

Performance Tests

Status:  NOT STARTED

Tasks:

- Load testing (100+ concurrent users)
 - Database query optimization
 - API response time testing
 - Memory usage monitoring
-

Deployment (0% Complete)

Production Setup

Status:  NOT STARTED

Tasks:

- Choose hosting platform (Heroku/Railway/DigitalOcean)
 - Set up production database (PostgreSQL)
 - Configure environment variables
 - Set up SSL certificate
 - Configure domain name
 - Set up monitoring (logs, errors)
 - Create backup strategy
-

CURRENT DATABASE STATUS

Data Collected So Far

Total Posts in Database: 908+ posts

- RSS Feeds: 733 posts
- Reddit: 175 posts
- News API: ~500-800 posts (pending test completion)
- Twitter: 0 posts (not integrated yet)

Geographic Coverage:

- Continents: 6 (all major continents)
- Countries: 40+ represented
- Cities: Data available but not extracted yet

Data Quality:

- Posts have: text, source, timestamp 
 - Posts have: country (most) 
 - Posts missing: emotion classification 
 - Posts missing: city/continent (some) 
-

TECHNICAL ISSUES & SOLUTIONS

Issue 1: spaCy Installation Failed RESOLVED

Problem: spaCy requires C++ Build Tools, incompatible with Python 3.14

Solution: Decided to use simpler regex + dictionary approach for location extraction

Impact: No impact on functionality, actually faster and simpler

Issue 2: Some RSS Feeds Invalid RESOLVED

Problem: 5 out of 20 RSS feeds returned errors or were invalid

Solution: Added error handling, still collecting from 15 working feeds (75% success rate)

Impact: Minor, still get 700+ posts from RSS alone

Issue 3: API Rate Limits MONITORING

Problem: NewsData.io has 100 requests/day limit

Solution:

- Running every 60 minutes (not 5 minutes)
- Collecting from 32 countries = 32 requests per run
- Can run ~3 times per day safely

Impact: Acceptable for MVP, will need premium plan for production

Issue 4: Emotion Classification Not Implemented PENDING

Problem: Currently no emotion analysis, only VADER sentiment

Solution: Will implement keyword-based emotion classifier (recommended approach in requirements doc)

Timeline: 1-2 days to implement



METRICS & KPIs

Data Collection Performance

Metric	Current	Target	Status
Posts/Hour	~900	2,000	● 45%
Sources Active	3/4	4/4	● 75%
Countries Covered	40+	45+	● 89%
Success Rate	85%	90%	● 94%
Update Frequency	60 min	60 min	● 100%

System Performance

Metric	Current	Target	Status
API Response Time	<100ms	<500ms	● Excellent
Database Size	~2MB	<100MB	● Excellent
Uptime	100%	99%	● Excellent

NEXT IMMEDIATE STEPS (Priority Order)

This Week (High Priority)

1. Complete News Collector Test (30 minutes)

- Run and verify 32-country collection
- Check database for new posts
- Validate data quality

2. Create Emotion Analyzer (1-2 days)

- Implement keyword-based classifier
- Test on existing 900+ posts
- Update database with emotions
- Validate accuracy

3. Create Location Extractor (1-2 days)

- Build city/country dictionary
- Implement regex patterns
- Test on existing posts
- Update database with locations

4. Create Data Aggregator (1 day)

- Aggregate by country
- Aggregate by continent
- Calculate emotion percentages
- Save to aggregated_emotions table

5. Build Frontend HTML/CSS (2-3 days)

- Create complete structure
- Style with emotion colors
- Make responsive

Next Week (Medium Priority)

6. Implement Plotly.js Map (2 days)

- Interactive world map
- Color by emotion
- Click handlers

7. Build Dashboard Components (2 days)

- Global emotion meter
- Statistics cards
- Trend charts

8. Integrate Twitter API (2-3 days)

- Get API access
- Implement collector
- Test and validate

9. Complete Background Scheduler (1 day)

- Integrate all collectors
- Add processing pipeline
- Test full automation

10. Testing & Bug Fixes (2-3 days)

- Test all features
- Fix bugs
- Performance optimization

TEAM RESPONSIBILITIES

Current Active Members

Victor (Lead)

-  Backend architecture
-  Database design
-  API endpoints
-  Integration coordination
-  Deployment

Team Member 2 (Backend)

-  RSS collector
-  News API collector
-  Emotion analyzer

Team Member 3 (Backend)

-  Reddit collector
-  Twitter collector
-  Location extractor

Team Members 4-5 (Frontend)

-  HTML/CSS design
-  JavaScript components
-  Map visualization

Team Member 6 (Data)

-  Emotion model validation
-  Data quality checks
-  Testing

Team Member 7 (DevOps)

-  Testing
-  Deployment setup
-  Documentation

TIMELINE & MILESTONES

Week 1 (Oct 18-24) - Foundation COMPLETE

-  Project setup
-  Database design
-  API framework

- 3 data collectors working

Week 2 (Oct 25-31) - Intelligence IN PROGRESS

- Emotion analyzer
- Location extractor
- Data aggregator
- Frontend structure

Week 3 (Nov 1-7) - Visualization PENDING

- Interactive map
 - Dashboard
 - Testing
 - MVP launch
-

LESSONS LEARNED

What Worked Well

1. **Using UV for package management** - Fast and reliable
2. **Starting with simple collectors** - RSS and Reddit easy wins
3. **Database-first approach** - Clear data structure from start
4. **Team collaboration** - API collectors done in parallel
5. **Git workflow** - Easy to track progress and collaborate

Challenges Faced

1. **Python 3.14 compatibility** - Some packages not ready
2. **API rate limits** - Had to reduce update frequency
3. **Location extraction complexity** - Decided to simplify approach
4. **RSS feed reliability** - Some feeds outdated/broken

Improvements for Next Sprint

1. Test all collectors before integration
 2. Document API requirements clearly
 3. Set up automated testing early
 4. Create demo data for frontend development
 5. Regular team sync meetings
-

BLOCKERS & NEEDS

Current Blockers

1. **Twitter API Access**  HIGH PRIORITY
 - Status: Waiting for approval
 - Impact: Missing real-time social data

- Workaround: Using Reddit as substitute
- Action: Team member 3 following up

2. Frontend Resources MEDIUM PRIORITY

- Status: Need designers to start
- Impact: MVP timeline at risk
- Action: Team members 4-5 to begin this week

Resource Needs

1. API Credits

- NewsData.io: Currently on free tier (100/day)
- Recommendation: Upgrade to paid plan (\$20/month) for production

2. Hosting

- Current: Local development
- Need: Production hosting (\$5-20/month)
- Options: Heroku, Railway.app, DigitalOcean

3. Domain Name

- Suggestion: pulsenet.io or emotionmap.io
- Cost: ~\$12/year

SUCCESS CRITERIA TRACKING

MVP Launch Checklist (Target: Nov 7)

Backend (70% Complete)

- Database operational
- 3+ data sources integrated
- Emotion analysis working
- Location extraction working
- Aggregation working
- API endpoints returning real data
- Background scheduler running

Frontend (0% Complete)

- Beautiful, responsive design
- Interactive map working
- Dashboard showing stats
- Search functionality
- Time filter/replay
- Mobile optimized

Quality (0% Complete)

- 70%+ emotion accuracy
- 40+ countries mapped
- <3 second page load
- 99% uptime for 24 hours
- No critical bugs

POST-MVP ROADMAP

Version 1.1 (Nov-Dec 2025)

- Twitter integration complete
- Real-time updates (WebSocket)
- User accounts
- Custom alerts
- Export reports

Version 1.2 (Jan 2026)

- Mobile app (React Native)
- Advanced AI insights
- Historical data analysis
- Predictive trends

Version 2.0 (Q1 2026)

- Business dashboard
- API for developers
- White-label solution
- Premium features

DAILY LOG

October 18, 2025

- Installed Python, UV, Git
- Created project structure
- Set up virtual environment
- Initialized database
- Created Flask app
- Pushed to GitHub

October 19, 2025

- Built API endpoints
- Created database manager
- Integrated RSS collector (733 posts!)
- Integrated Reddit collector (175 posts!)
- Set up configuration system

October 20, 2025 (Today)

- Fixed News API collector
- Reviewed project requirements (PulseNet PDF)
- Updated project vision to 5 emotions
- Created comprehensive documentation

- Planning emotion analyzer approach
-

RECOMMENDATIONS

Technical Decisions

1. Emotion Analysis RECOMMENDED

- **Use:** Keyword-based + VADER combination
- **Why:** Simple, fast, accurate enough for MVP
- **Alternative:** Pre-trained model (more complex, slower)

2. Location Extraction RECOMMENDED

- **Use:** Regex + Dictionary + Source mapping
- **Why:** Fast, reliable, easy to debug
- **Alternative:** spaCy NER (complex, installation issues)

3. Database CURRENT APPROACH GOOD

- **Keep:** SQLite for MVP
- **Upgrade to:** PostgreSQL for production
- **When:** After reaching 100K+ posts

4. Frontend Framework RECOMMENDED

- **Use:** Vanilla JavaScript + Plotly.js
- **Why:** No build tools needed, fast development
- **Alternative:** React (overkill for MVP)

Process Improvements

1. **Daily Standups** - 15 min team sync
 2. **Code Reviews** - Before merging to main
 3. **Testing First** - Write tests before coding new features
 4. **Documentation** - Update docs as you build
 5. **Demo Often** - Show progress to stakeholders weekly
-

WINS & CELEBRATIONS

Major Achievements

1.  **Database Design** - Clean schema, good performance
2.  **3 Collectors Working** - 900+ posts collected
3.  **API Framework Complete** - All endpoints working
4.  **GitHub Setup** - Easy collaboration
5.  **Clear Documentation** - Easy for team to understand

Team Kudos 🌟

- **Victor:** Excellent project leadership and architecture
 - **Backend Team:** Great API collector implementations
 - **Everyone:** Strong collaboration and communication
-

RESOURCES & LINKS

Project Links

- **Repository:** <https://github.com/vixtor-e86/emotion-map-project>
- **Local Server:** <http://localhost:5000>
- **API Docs:** (To be created)

External APIs

- **NewsData.io:** <https://newsdata.io/>
- **Reddit PRAW:** <https://praw.readthedocs.io/>
- **Twitter API:** <https://developer.twitter.com/>
- **Feedparser:** <https://feedparser.readthedocs.io/>

Learning Resources

- **Plotly Maps:** <https://plotly.com/javascript/choropleth-maps/>
 - **VADER Sentiment:** <https://github.com/cjhutto/vaderSentiment>
 - **Flask Tutorial:** <https://flask.palletsprojects.com/>
-

📞 CONTACT & SUPPORT

Project Lead: Victor (@vixtor-e86)

GitHub Issues: [Repository Issues Page]

Team Chat: [Add communication channel]

Weekly Meeting: [Add meeting schedule]

Last Updated: October 20, 2025 - 4:30 PM

Next Update: October 21, 2025

Document Maintained By: Victor

🚀 READY TO LAUNCH

Current Status: 55% Complete

MVP Target Date: November 7, 2025

Days Remaining: 18 days

Confidence Level: HIGH 

We're on track! Keep pushing! 