

คู่มือสถาปัตยกรรมและการปฏิบัติการระบบ TrendSiam

คู่มือฉบับสมบูรณ์สำหรับนักพัฒนาและผู้ดูแลระบบ



เวอร์ชัน 1.0

วันที่จัดทำ: 20 สิงหาคม 2568

ระบบรวบรวมและวิเคราะห์ข่าวแนวโน้มประเทศไทย
Thai News Trend Aggregation & Analysis System

สารบัญ

1. ภาพรวมสถาปัตยกรรมระบบ	หน้า 3
2. การไหลของข้อมูล	หน้า 4
3. ความปลอดภัยของระบบ	หน้า 5
4. การใช้งาน Supabase	หน้า 6
5. คู่มือการปฏิบัติการ	หน้า 7
6. แนวทางการทดสอบ	หน้า 8
ภาคผนวก ก: บัญชีรายชื่อไฟล์	หน้า 9
ภาคผนวก ข: รายการ API และเส้นทาง	หน้า 10+

หมายเหตุ

- เอกสารนี้จัดทำขึ้นเพื่อใช้เป็นคู่มืออ้างอิงสำหรับนักพัฒนาและผู้ดูแลระบบ
- ข้อมูลในเอกสารอาจมีการเปลี่ยนแปลงตามการพัฒนาาระบบ
- สำหรับข้อมูลล่าสุด กรุณาตรวจสอบจากเอกสารออนไลน์หรือติดต่อทีมพัฒนา

1. ภาพรวมสถาปัตยกรรมระบบ

1.1 ภาพรวมระบบ TrendSiam

ระบบ TrendSiam เป็นแพลตฟอร์มรวบรวมและวิเคราะห์ข่าวแนวโน้มจากแหล่งข้อมูลต่างๆ โดยเฉพาะจาก YouTube และสื่อออนไลน์ในประเทศไทย ระบบใช้เทคโนโลยี AI ในการจัดหมวดหมู่ สร้างสรุป และสร้างภาพประกอบอัตโนมัติ

1.2 องค์ประกอบหลักของระบบ

Frontend (Next.js)

- เว็บแอปพลิเคชันสำหรับแสดงผลข้อมูลแนวโน้ม
- รายงาน PDF รายสัปดาห์
- หน้าจอสําหรับผู้ดูแลระบบ
- API endpoints สำหรับการเชื่อมต่อข้อมูล

Backend Processing (Python)

- ระบบดึงข้อมูลจาก YouTube API
- การประมวลผลและจัดหมวดหมู่ด้วย AI
- การสร้างสรุปภาษาไทยและอังกฤษ
- การสร้างภาพประกอบด้วย AI
- การคำนวณคะแนนความนิยม

ฐานข้อมูล (Supabase)

- PostgreSQL database สำหรับเก็บข้อมูลข่าว
- Authentication และ authorization
- Storage สำหรับไฟล์ภาพ
- Real-time subscriptions
- Row Level Security (RLS) policies

1.3 เทคโนโลยีที่ใช้

Frontend

- Next.js 14
- React 18
- TypeScript
- Tailwind CSS

Backend

- Python 3.11+
- OpenAI API
- YouTube Data API
- Supabase Python Client

- React-PDF
- Zustand (State Management)

- Pillow (Image Processing)
- BeautifulSoup4

1.4 การติดตั้งและใช้งาน

ระบบสามารถติดตั้งและใช้งานได้หลายรูปแบบ:

- Development: Local development server
- Staging: Vercel หรือ cloud platform อื่นๆ
- Production: Vercel + Supabase Cloud
- Self-hosted: Docker containers

2. การไหลของข้อมูล

2.1 การรวบรวมข้อมูล

ขั้นตอนที่ 1: ดึงข้อมูลจาก YouTube

- ใช้ YouTube Data API v3 ดึงข้อมูลวิดีโอยอดนิยม
- กรองข้อมูลตามภูมิภาค (ประเทศไทย)
- ดึงข้อมูล metadata: ชื่อ, คำอธิบาย, จำนวนวิว, โไลค์, คอมเมนต์
- บันทึกข้อมูลดิบลงฐานข้อมูล

ขั้นตอนที่ 2: การประมวลผลด้วย AI

- ส่งข้อมูลไปยัง OpenAI API สำหรับการวิเคราะห์
- จัดหมวดหมู่เนื้อหา (บันเทิง, ข่าว, กีฬา, เทคโนโลยี, ฯลฯ)
- สร้างสรุปภาษาไทยและอังกฤษ
- สกัดคำสำคัญและแท็ก
- คำนวณคะแนนความเกี่ยวข้องและความนิยม

2.2 Pipeline การประมวลผล

YouTube API ' Raw Data ' AI Processing ' Categorization ' Summary Generation ' Image Generation ' Score Calculation ' Database Storage ' Frontend Display

รายละเอียดแต่ละขั้นตอน:

- Raw Data Collection: ดึงข้อมูลจาก API และบันทึกในรูปแบบดิบ
- Content Analysis: วิเคราะห์เนื้อหาด้วย AI เพื่อเข้าใจบริบท
- Categorization: จัดหมวดหมู่ตามประเภทเนื้อหา
- Summary Generation: สร้างสรุปที่เข้าใจง่าย
- Image Generation: สร้างภาพประกอบด้วย DALL-E
- Quality Scoring: คำนวณคะแนนความนิยมและคุณภาพ

2.3 โครงสร้างฐานข้อมูล

ตารางหลักในระบบ:

- news_trends: ข้อมูลข่าวและแนวโน้มหลัก
- weekly_report_snapshots: ข้อมูลรายงานรายสัปดาห์

- ai_generated_images: ภาพที่สร้างด้วย AI

2.4 การอัปเดตแบบ Real-time

ระบบใช้ Supabase Real-time สำหรับการอัปเดตข้อมูลแบบทันที

- การแจ้งเตือนเมื่อมีข้อมูลใหม่
- การอัปเดตสถิติแบบ live
- การซิงค์ข้อมูลระหว่างหน้าต่างๆ
- การแจ้งเตือนสถานะการประมวลผล

3. ความปลอดภัยของระบบ

3.1 ภาพรวมความปลอดภัย

ระบบ TrendSiam ได้รับการออกแบบด้วยหลักการความปลอดภัยแบบ Defense in Depth โดยมีการป้องกันหลายชั้นตั้งแต่ระดับเครือข่าย แอปพลิเคชัน จนถึงฐานข้อมูล

3.2 ความปลอดภัยของเว็บแอปพลิเคชัน

Content Security Policy (CSP)

- ป้องกันการโจมตีแบบ XSS (Cross-Site Scripting)
- จำกัดแหล่งที่มาของ script, style, และ image
- บล็อกการโหลด inline script ที่ไม่ปลอดภัย
- รายงานการละเมิดนโยบายไปยัง monitoring system

HTTP Security Headers

- X-Frame-Options: ป้องกัน clickjacking
- X-Content-Type-Options: ป้องกัน MIME type sniffing
- Referrer-Policy: ควบคุมการส่ง referrer information
- Strict-Transport-Security: บังคับใช้ HTTPS

Input Validation และ Sanitization

- ตรวจสอบและทำความสะอาดข้อมูลทุก API endpoint
- ใช้ Zod schema สำหรับ type validation
- Sanitize ข้อมูลก่อนแสดงผลใน PDF
- ป้องกัน SQL injection ด้วย parameterized queries

3.3 ความปลอดภัยของ API

Rate Limiting

- จำกัดจำนวน request ต่อ IP address
- ป้องกันการโจมตีแบบ DDoS
- ใช้ sliding window algorithm
- แจ้งเตือนเมื่อมีการใช้งานผิดปกติ

Authentication และ Authorization

- ใช้ Supabase Auth สำหรับการจัดการผู้ใช้
- JWT tokens สำหรับ session management
- Role-based access control (RBAC)
- Multi-factor authentication (MFA) สำหรับ admin

3.4 การจัดการข้อผิดพลาด

หลักการสำคัญในการจัดการข้อผิดพลาด

- ไม่เปิดเผยข้อมูลระบบภายในในข้อความ error
- Log ข้อผิดพลาดแบบละเอียดใน server-side เท่านั้น
- ส่งข้อความ error ที่เป็นมิตรกับผู้ใช้
- ไม่ log ข้อมูลที่เป็นความลับ (API keys, passwords)
- ใช้ structured logging สำหรับการวิเคราะห์

3.5 การป้องกัน SSRF และ Path Traversal

- ตรวจสอบ URL ก่อนทำ HTTP requests
- บล็อก private IP ranges และ localhost
- Whitelist allowed domains สำหรับ external requests
- Validate file paths และป้องกัน directory traversal
- ใช้ sandbox environment สำหรับการประมวลผลไฟล์

4. การใช้งาน Supabase

4.1 ภาพรวม Supabase

Supabase เป็น Backend-as-a-Service ที่ให้บริการ PostgreSQL database, Authentication, Storage, และ Real-time subscriptions ระบบ TrendSiam ใช้ Supabase เป็นแกนหลักสำหรับการจัดเก็บและจัดการข้อมูล

4.2 โครงสร้างฐานข้อมูล

ตารางหลัก

news_trends

เก็บข้อมูลข่าวและแนวโน้มหลัก รวมถึง metadata, คะแนน, และ AI analysis

weekly_report_snapshots

เก็บข้อมูล snapshot สำหรับรายงานรายสัปดาห์ ใช้สำหรับสร้าง PDF

ai_generated_images

เก็บข้อมูลภาพที่สร้างด้วย AI รวมถึง prompts และ metadata

Views และ Functions

- weekly_report_public_v: View สำหรับข้อมูลรายงานสาธารณะ
- Various helper functions สำหรับการคำนวณและการจัดการข้อมูล

4.3 Row Level Security (RLS)

การตั้งค่า RLS policies สำหรับความปลอดภัย

- Public Read Access: อนุญาตให้ anon role อ่านข้อมูลที่เผยแพร่แล้ว
- Admin Write Access: เฉพาะ service role เท่านั้นที่เขียนข้อมูลได้
- Status-based Filtering: กรองข้อมูลตาม status (published, draft, etc.)

4.4 Authentication และ Authorization

Client-side Authentication

- ใช้ anon key สำหรับการเข้าถึงข้อมูลสาธารณะ
- JWT tokens สำหรับ authenticated users
- Session management ผ่าน Supabase Auth

Server-side Operations

- ใช้ service role key สำหรับ backend operations
- CLI scripts และ cron jobs
- Admin operations และ data processing

4.5 Storage Configuration

การจัดการไฟล์และภาพ:

- ai-generated-images: Bucket สำหรับภาพที่สร้างด้วย AI
- Public Access: ภาพสามารถเข้าถึงได้โดยไม่ต้อง authentication
- CDN Integration: ใช้ Supabase CDN สำหรับการแจกจ่ายไฟล์

4.6 แนวทางปฏิบัติที่ดี

- แยก environment variables ระหว่าง development และ production
- ใช้ connection pooling สำหรับ high-traffic applications
- Monitor database performance และ query optimization
- Backup ข้อมูลสำคัญเป็นประจำ
- ตรวจสอบ RLS policies เป็นระยะ
- ใช้ prepared statements เพื่อป้องกัน SQL injection

5. คู่มือการปฏิบัติการ

5.1 การติดตั้งและใช้งาน

การติดตั้งเบื้องต้น

```
# Clone repository
git clone https://github.com/your-org/trendsiam.git
cd trendsiam

# Install dependencies
cd frontend
npm install

# Setup environment
cp .env.example .env.local
# Edit .env.local with your Supabase credentials
```

การรันระบบ

```
# Development server
npm run dev

# Production build
npm run build
npm run start

# Test weekly PDF generation
curl -o weekly.pdf "http://localhost:3000/api/weekly/pdf"
```

5.2 การประมวลผลข้อมูล

การดึงข้อมูลข่าวใหม่

```
# 6 I!9% 2#%H2*8
python summarize_all_v2.py --20

# 6 I!9%A @ G!
python summarize_all_v2.py
```

```
# #!python3
python scripts/check_pipeline_health.py
```

การสร้างรายงานรายสัปดาห์

```
# #!bash
cd frontend
npm run snapshot:build:publish

# #!bash
npm run snapshot:test:selection

# #!bash
curl http://localhost:3000/api/weekly/diagnostics
```

5.3 การตรวจสอบสุขภาพระบบ

Health Check Endpoints

- /api/health - สถานะระบบโดยรวม
- /api/health/db - การเชื่อมต่อฐานข้อมูล
- /api/weekly/diagnostics - สถานะรายงานรายสัปดาห์
- /api/env-check - ตรวจสอบ environment variables

การตรวจสอบข้อมูล

```
# #!python3
python test_pipeline_diagnostics.py

# #!python3
python scripts/verify_images.py

# #!python3
python scripts/check_score_details.py
```

5.4 การแก้ไขปัญหาเบื้องต้น

ปัญหา: Weekly report แสดง "No snapshots available"

- ตรวจสอบ /api/weekly/diagnostics
- รัน npm run snapshot:build:publish
- ตรวจสอบ environment variables

ปัญหา: AI image generation ล้มเหลว

- ตรวจสอบ OpenAI API key
- ตรวจสอบ quota และ rate limits
- รัน python scripts/verify_images.py

ปัญหา: Database connection error

- ตรวจสอบ Supabase credentials
- ตรวจสอบ network connectivity
- ตรวจสอบ RLS policies

6. แนวทางการทดสอบ

6.1 ภาพรวมการทดสอบ

ระบบ TrendSiam มีการทดสอบหลายระดับเพื่อให้มั่นใจในคุณภาพและความเสถียร ตั้งแต่ unit tests, integration tests, จนถึง end-to-end tests

6.2 การทดสอบ Frontend

การทดสอบพื้นฐาน

```
# Type checking
npm run type-check

# Linting
npm run lint

# Build test
npm run build

# Security scan
npm run scan:build:secrets
```

การทดสอบ Weekly Report System

```
# *- 2# 1 I-I9%
npm run snapshot:test:count

# *- 2#@%7- I-I9%
npm run snapshot:test:selection

# *- 2# 1 I-I9%
npm run snapshot:test
```

การทดสอบ PDF Generation

```
# PDF overlap fix
npm run test:pdf-overlap

# Weekly PDF
curl -o test-weekly.pdf "http://localhost:3000/api/weekly/pdf"
```

```
# # 5H*#I2
# @ 4 D %LA%0 #' *- 'H2D!H!5 I- '2! 1 I-
```

6.3 การทดสอบ Backend

การทดสอบ Data Pipeline

```
# *- 2# 6 I!9%
python test_pipeline_diagnostics.py

# AI processing
python test_ai_image_generation.py

# Categorization
python test_categories_pipeline.py

# Keyword extraction
python -m pytest tests/test_keyword_extraction.py
```

การทดสอบ Security

```
# Crypto security
python -m pytest tests/test_crypto_security.py

# Image URL protection
python -m pytest tests/test_image_url_protection.py

# Security audit
python security_audit.py
```

6.4 การทดสอบแบบ Integration

End-to-End Testing

```
# E2E fixes
python validate_e2e_fixes.py

# Acceptance
python acceptance_tests.py
```

```
# compatibility
python test_compat_scores.py
```

การทดสอบ API Endpoints

- GET /api/health - ควรได้ status 200
- GET /api/weekly/diagnostics - ตรวจสอบข้อมูล snapshot
- GET /api/weekly/pdf - ควรได้ไฟล์ PDF
- POST /api/admin/revalidate - ต้องมี secret header

6.5 การทดสอบประสิทธิภาพ

แนวทางการทดสอบประสิทธิภาพ:

- Load testing สำหรับ API endpoints
- Database query performance monitoring
- PDF generation time measurement
- Memory usage profiling
- Rate limiting effectiveness
- CDN cache hit rates

6.6 Checklist การทดสอบก่อน Deploy

Type checking ผ่าน (npm run type-check)
Build สำเร็จ (npm run build)
Security scan ผ่าน (npm run scan:build:secrets)
Weekly report ทำงานได้ (npm run snapshot:test:selection)
PDF generation ทำงานได้
Backend tests ผ่าน
API endpoints ตอบสนองปกติ
Database connection ใช้งานได้
Environment variables ครบถ้วน

ภาคผนวก ก: บัญชีรายชื่อไฟล์

บัญชีรายชื่อไฟล์ทั้งหมดในระบบ TrendSiam โดยเฉพาะไฟล์ Python ที่เป็นส่วนสำคัญ ของการประมวลผลข้อมูลและ AI processing

ก.1 ไฟล์ Python หลัก

Root Directory

acceptance_tests.py

ชุดทดสอบการยอมรับสำหรับตรวจสอบฟังก์ชันหลัก

ฟังก์ชันหลัก: test_a_force_refresh, test_b_verify_api_parity, test_c_manual_ui_check, test_d_dry_run_safety, run_all_acceptance_tests...

การใช้งาน: Test suite, Database operations

ขนาดไฟล์: 10.4 KB | เส้นทาง: acceptance_tests.py

ai_image_generator.py

การประมวลผลภาพและ AI image generation สำหรับ ai_image_generator

ฟังก์ชันหลัก: get_precise_score, sanitize_prompt_text, generate_safe_fallback_prompt, generate_ai_images_for_top3_news

การใช้งาน: CLI script

ขนาดไฟล์: 45.4 KB | เส้นทาง: ai_image_generator.py

ai_image_generator_v2.py

ระบบสร้างภาพประกอบด้วย AI สำหรับข่าวและเนื้อหา

ฟังก์ชันหลัก: generate_story_id, parse_publish_time, get_precise_score, validate_image_file, sanitize_prompt_text...

การใช้งาน: CLI script

ขนาดไฟล์: 25.6 KB | เส้นทาง: ai_image_generator_v2.py

ai_image_supabase_generator.py

การประมวลผลภาพและ AI image generation สำหรับ ai_image_supabase_generator

การใช้งาน: Database operations

ขนาดไฟล์: 8.1 KB | เส้นทาง: ai_image_supabase_generator.py

app.py

TrendSiam - Thai Daily News Summary Web App

A simple Streamlit application that displays summarized Thai trending news from YouTube videos with AI-generated Thai summaries. (โมดูล Python สำหรับ app)

ฟังก์ชันหลัก: apply_dynamic_theme, apply_radio_theme_css, is_dev_mode, validate_youtube_video_id, create_secure_youtube_url...

การใช้งาน: CLI script

📁archive\app_original.py

TrendSiam - Thai Daily News Summary Web App

A simple Streamlit application that displays summarized Thai trending news from YouTube videos with AI-generated Thai summaries. (โหลด Python สำหรับ archive\app_original)

ฟังก์ชันหลัก: apply_dynamic_theme, apply_radio_theme_css, is_dev_mode, validate_youtube_video_id, create_secure_youtube_url...

การใช้งาน: CLI script

ขนาดไฟล์: 212 KB | เส้นทาง: archive\app_original.py

📁clean_trending_data.py

เครื่องมือทำความสะอาดและปรับปรุงข้อมูลแนวโน้ม

ฟังก์ชันหลัก: is_fake_video_id, is_fake_channel, is_fake_title, has_suspicious_view_count, is_fake_entry...

การใช้งาน: CLI script

ขนาดไฟล์: 8 KB | เส้นทาง: clean_trending_data.py

📁cleanup_backup_20250728_200913\fix_dev_password.py

Script to fix the developer password security vulnerability in app.py (โหลด Python สำหรับ cleanup_backup_20250728_200913\fix_dev_password)

ฟังก์ชันหลัก: fix_dev_password

การใช้งาน: CLI script

ขนาดไฟล์: 1.9 KB | เส้นทาง: cleanup_backup_20250728_200913\fix_dev_password.py

📁config\constants.py

การตั้งค่าและกำหนดค่าสำหรับ config\constants

ฟังก์ชันหลัก: get_developer_password, get_feature_flag

ขนาดไฟล์: 6 KB | เส้นทาง: config\constants.py

📁config_openai.py

การตั้งค่าและกำหนดค่าสำหรับ config_openai

การใช้งาน: CLI script

ขนาดไฟล์: 2.5 KB | เส้นทาง: config_openai.py

📁core\api_security.py

API endpoints และการจัดการ requests สำหรับ core\api_security

ฟังก์ชันหลัก: rate_limit, secure_api_call

ขนาดไฟล์: 14.7 KB | เส้นทาง: core\api_security.py

📁core\config.py

การตั้งค่าและกำหนดค่าสำหรับ core\config

ฟังก์ชันหลัก: get_config, reload_config, get_openai_config

core\constants.py

Constants and configuration values for TrendSiam application.

This file centralizes all magic numbers and hardcoded values to improve maintainability. (โมดูล Python สำหรับ core\constants)

ขนาดไฟล์: 3.3 KB | เส้นทาง: core\constants.py

core\crypto_security.py

Cryptography Security Wrapper for TrendSiam

This module provides secure cryptographic operations and prevents usage of vulnerable APIs. All cryptographic operations should go through this module.

Security measures:

- Blocks RSA PKCS1v15 padding for decryption (forces OAEP)
- Validates PKCS#12 certificate/key matching
- Provides secure defaults for all operations (โมดูล Python สำหรับ core\crypto_security)

ฟังก์ชันหลัก: detect_insecure_crypto_usage

การใช้งาน: CLI script

ขนาดไฟล์: 10.1 KB | เส้นทาง: core\crypto_security.py

core\env_config.py

การตั้งค่าและกำหนดค่าสำหรับ core\env_config

ฟังก์ชันหลัก: load_config, get_config, create_template

การใช้งาน: CLI script, Database operations

ขนาดไฟล์: 12.7 KB | เส้นทาง: core\env_config.py

core__init__.py

TrendSiam Core Security Module

This package provides comprehensive security, privacy, and compliance management for the TrendSiam application including:

- Secure configuration management with environment variable protection
- Input validation and sanitization with XSS/injection prevention
- Comprehensive logging with credential sanitization
- Secure subprocess execution with command injection prevention
- API security with rate limiting and request validation
- Legal compliance for YouTube, OpenAI, and DALL-E APIs
- Privacy protection and GDPR compliance
- Ethical AI usage validation

Security Features:

Environment variable protection

API key validation and secure storage

Input sanitization and validation

Command injection prevention
Rate limiting and abuse prevention
Legal compliance checking
Privacy-compliant data handling
Comprehensive security logging
Error handling with information disclosure prevention (โหลด Python สำหรับ core__init__)

ฟังก์ชันหลัก: initialize_security, run_security_audit
ขนาดไฟล์: 8.5 KB | เส้นทาง: core__init__.py

core\legal_compliance.py
Legal Compliance Module for TrendSiam

This module ensures compliance with Terms of Service, privacy regulations, and responsible data usage for all integrated platforms and services. (โหลด Python สำหรับ core\legal_compliance)

ฟังก์ชันหลัก: get_compliance_manager, ensure_youtube_compliance, ensure_openai_compliance, ensure_dalle_compliance
ขนาดไฟล์: 17.1 KB | เส้นทาง: core\legal_compliance.py

core\logging_config.py
การตั้งค่าและกำหนดค่าสำหรับ core\logging_config

ฟังก์ชันหลัก: performance_monitor, error_handler, create_module_logger, setup_logging, get_app_logger
ขนาดไฟล์: 11.5 KB | เส้นทาง: core\logging_config.py

core\secure_subprocess.py
Secure Subprocess Wrapper for TrendSiam

This module provides secure subprocess execution with input validation, command injection prevention, and comprehensive logging. (โหลด Python สำหรับ core\secure_subprocess)

ฟังก์ชันหลัก: secure_subprocess_run, secure_ytdlp_call
ขนาดไฟล์: 11.1 KB | เส้นทาง: core\secure_subprocess.py

core\storage_config.py
การตั้งค่าและกำหนดค่าสำหรับ core\storage_config

ฟังก์ชันหลัก: get_storage_config, get_bucket_name, get_image_public_url, validate_storage
การใช้งาน: CLI script, Database operations
ขนาดไฟล์: 4.7 KB | เส้นทาง: core\storage_config.py

core\validators.py

Input Validation and Security Module for TrendSiam

This module provides comprehensive input validation, sanitization, and security checks for all external data sources. (โมดูล Python สำหรับ core\validators)

ฟังก์ชันหลัก: validate_news_data, safe_get_env_var

ขนาดไฟล์: 13.5 KB | เส้นทาง: core\validators.py

debug_second_story_image.py

การประมวลผลภาพและ AI image generation สำหรับ debug_second_story_image

ฟังก์ชันหลัก: check_supabase_data, check_frontend_api, main

การใช้งาน: CLI script, Database operations

ขนาดไฟล์: 6.5 KB | เส้นทาง: debug_second_story_image.py

diagnose_supabase.py

Diagnostic script to test Supabase connection and identify issues (โมดูล Python สำหรับ diagnose_supabase)

การใช้งาน: Database operations

ขนาดไฟล์: 7.1 KB | เส้นทาง: diagnose_supabase.py

fix_ai_images_for_existing.py

การประมวลผลภาพและ AI image generation สำหรับ fix_ai_images_for_existing

ฟังก์ชันหลัก: generate_images_for_top3

การใช้งาน: CLI script, Database operations

ขนาดไฟล์: 3.9 KB | เส้นทาง: fix_ai_images_for_existing.py

fix_env_DANGEROUS_DELETE_ME.py

โมดูล Python: fix_env_DANGEROUS_DELETE_ME

ขนาดไฟล์: 139 B | เส้นทาง: fix_env_DANGEROUS_DELETE_ME.py

force_regenerate_images.py

การประมวลผลภาพและ AI image generation สำหรับ force_regenerate_images

ฟังก์ชันหลัก: force_regenerate_images, main

การใช้งาน: CLI script

ขนาดไฟล์: 4.5 KB | เส้นทาง: force_regenerate_images.py

generate_html_pdf.py

TrendSiam HTML Report Generator with WeasyPrint

Purpose: Generate professional PDF reports using HTML/CSS template and WeasyPrint (โมดูล Python สำหรับ generate_html_pdf)

ฟังก์ชันหลัก: safe_int, safe_float, parse_view_count, get_precise_score, get_score_class...

การใช้งาน: CLI script

ขนาดไฟล์: 22.5 KB | เส้นทาง: generate_html_pdf.py

🔗 legacy_adapter.py

Legacy Adapter for TrendSiam V2 Compatibility

This module provides adapters to reuse legacy popularity scoring and category classification logic in the V2 ingestion system, ensuring 100% compatibility. (โมดูล Python สำหรับ legacy_adapter)

ขนาดไฟล์: 10.1 KB | เส้นทาง: legacy_adapter.py

🔗 mark_json_legacy.py

SECTION H - JSON Legacy Marking Script

This script marks JSON data files as legacy by adding .legacy suffix while preserving functionality for emergency use. (โมดูล Python สำหรับ mark_json_legacy)

ฟังก์ชันหลัก: mark_json_files_as_legacy

การใช้งาน: CLI script

ขนาดไฟล์: 2.2 KB | เส้นทาง: mark_json_legacy.py

🔗 popularity_scorer.py

ระบบคำนวณคะแนนความนิยมและการจัดอันดับ

ฟังก์ชันหลัก: score_video_popularity, add_popularity_scores, main

การใช้งาน: CLI script

ขนาดไฟล์: 37.3 KB | เส้นทาง: popularity_scorer.py

🔗 scripts\check_aux_fields_pipeline.py

Check auxiliary fields in the pipeline

Verify which fields are being computed and written (โมดูล Python สำหรับ scripts\check_aux_fields_pipeline)

ฟังก์ชันหลัก: check_pipeline_fields

การใช้งาน: CLI script, Database operations

ขนาดไฟล์: 2.2 KB | เส้นทาง: scripts\check_aux_fields_pipeline.py

🔗 scripts\check_pipeline_health.py

SECTION G - Pipeline Health Check Script

Reads system_meta('news_last_updated'), prints age in minutes, and top 1 story.

Exit non-zero if age > 180 minutes in production. (โมดูล Python สำหรับ scripts\check_pipeline_health)

ฟังก์ชันหลัก: get_supabase_client, get_last_updated, get_top_story, calculate_age_minutes, main

การใช้งาน: Database operations

ขนาดไฟล์: 3.8 KB | เส้นทาง: scripts\check_pipeline_health.py

🔗 scripts\check_score_details.py

Check score_details column and data in the database (โมดูล Python สำหรับ scripts\check_score_details)

การใช้งาน: Database operations

ขนาดไฟล์: 2.5 KB | เส้นทาง: scripts\check_score_details.py

scripts\find_unused_files.py

Find Unused Files in TrendSiam Repository

This script identifies potentially unused files that can be safely removed.
It generates a dry-run report before any deletion. (โมดูล Python สำหรับ scripts\find_unused_files)

ฟังก์ชันหลัก: main

การใช้งาน: CLI script, Database operations

ขนาดไฟล์: 16.3 KB | เส้นทาง: scripts\find_unused_files.py

scripts\fix_single_video_enrichment.py

Fix missing auxiliary fields for a single video
Specifically handles BNK48 and other edge cases (โมดูล Python สำหรับ scripts\fix_single_video_enrichment)

ฟังก์ชันหลัก: extract_keywords_multilingual, calculate_growth_rate_safe, extract_platform_mentions, generate_score_details, generate_ai_opinion...

การใช้งาน: CLI script, Database operations

ขนาดไฟล์: 10.6 KB | เส้นทาง: scripts\fix_single_video_enrichment.py

scripts\security_audit_comprehensive.py

Comprehensive Security Audit for TrendSiam Repository

This script performs a thorough security audit including:

1. Secret detection
2. Dependency vulnerabilities
3. Code security patterns
4. Configuration issues (โมดูล Python สำหรับ scripts\security_audit_comprehensive)

ฟังก์ชันหลัก: main

การใช้งาน: CLI script, Database operations

ขนาดไฟล์: 18.4 KB | เส้นทาง: scripts\security_audit_comprehensive.py

scripts\test_bnk48_fix.py

ไฟล์ทดสอบสำหรับ scripts\bnk48_fix

ฟังก์ชันหลัก: extract_keywords_old, extract_keywords_new, detect_platforms_old, detect_platforms_new, generate_opinion_old...

การใช้งาน: CLI script, Test suite

ขนาดไฟล์: 6.4 KB | เส้นทาง: scripts\test_bnk48_fix.py

scripts\verify_image_setup.py

การประมวลผลภาพและ AI image generation สำหรับ scripts\verify_image_setup

ฟังก์ชันหลัก: verify_storage_setup, verify_database_state, main

การใช้งาน: CLI script, Database operations

ขนาดไฟล์: 6.6 KB | เส้นทาง: scripts\verify_image_setup.py

scripts\verify_images.py

การประมวลผลภาพและ AI image generation สำหรับ scripts\verify_images

ฟังก์ชันหลัก: verify_supabase_connection, get_recent_items_without_images, verify_image_url, check_existing_images, regenerate_missing_images...

การใช้งาน: CLI script, Database operations

ขนาดไฟล์: 8.9 KB | เส้นทาง: scripts\verify_images.py

scripts\verify_image_urls.py

การประมวลผลภาพและ AI image generation สำหรับ scripts\verify_image_urls

ฟังก์ชันหลัก: verify_image_urls

การใช้งาน: CLI script, Database operations

ขนาดไฟล์: 4.2 KB | เส้นทาง: scripts\verify_image_urls.py

scripts\verify_score_details_pipeline.py

Verify score details are being generated and written correctly (โมดูล Python สำหรับ scripts\verify_score_details_pipeline)

การใช้งาน: Database operations

ขนาดไฟล์: 3.1 KB | เส้นทาง: scripts\verify_score_details_pipeline.py

security_audit.py

เครื่องมือตรวจสอบความปลอดภัยของระบบ

ฟังก์ชันหลัก: audit_frontend_security, audit_environment_usage, print_audit_results

การใช้งาน: CLI script, Database operations

ขนาดไฟล์: 7.7 KB | เส้นทาง: security_audit.py

setup_backend_env.py

Helper script to set up the backend .env file with proper Supabase credentials (โมดูล Python สำหรับ setup_backend_env)

การใช้งาน: Database operations

ขนาดไฟล์: 3.5 KB | เส้นทาง: setup_backend_env.py

setup_environment.py

สคริปต์ติดตั้งค่าสภาพแวดล้อมและ dependencies

ฟังก์ชันหลัก: create_env_file, test_with_fallback, check_existing_env, main

การใช้งาน: CLI script, Test suite, Database operations

ขนาดไฟล์: 5.3 KB | เส้นทาง: setup_environment.py

summarize_all.py

TrendSiam News Ingestion Pipeline with Idempotency and Image Persistence

This script implements a robust news ingestion pipeline with the following guarantees:

A) Idempotency without losing history:

- Two-layer model: stories (canonical) + snapshots (per-run/day)
- Never destroys historical data
- Re-runs create/update snapshots, preserving history

- Atomic writes and non-destructive DB upserts only

B) Image persistence and regeneration policy (Top-3 focus):

- Never deletes/overwrites valid existing images
- Generates new images only when missing/invalid
- Stable story_id-based image mapping
- Top-3 image validation with retry logic

C) Ordering and alignment:

- Deterministic Top-3 using popularity_score desc, publish_time desc, story_id
- story_id and rank included in all outputs for UI alignment
- Images never reordered independently of stories

D) UX/UI and caching safety:

- Frontend JSON includes story_id, rank, image_status, data_version
- Placeholder handling for pending images
- Cache busting with data_version timestamp

E) Reliability and logging:

- Structured logging for all operations
- Exit codes: 0 (success), 5 (partial), others (errors)
- Configurable retry logic with exponential backoff
- Dry-run mode for testing (โมดูล Python สำหรับ summarize_all)

ฟังก์ชันหลัก: get_precise_score, generate_story_id, parse_publish_time, validate_image_file, determine_top3_ordering...

การใช้งาน: [CLI script](#), [Database operations](#)

ขนาดไฟล์: 85.7 KB | เส้นทาง: summarize_all.py

➤ summarize_all_v2.py

สคริปต์หลักสำหรับดึงข้อมูลและสร้างสรุปข่าวจาก YouTube API พร้อมการประมวลผล AI

ฟังก์ชันหลัก: generate_story_id, parse_publish_time, get_precise_score, determine_top3_ordering, validate_image_file...

การใช้งาน: [CLI script](#), [Database operations](#)

ขนาดไฟล์: 87.5 KB | เส้นทาง: summarize_all_v2.py

➤ summarize_all_v3_supabase_only.py

เวอร์ชันที่ใช้ Supabase เป็นหลักสำหรับการจัดเก็บและประมวลผลข้อมูล

ฟังก์ชันหลัก: parse_args, main

การใช้งาน: [Database operations](#)

ขนาดไฟล์: 22 KB | เส้นทาง: summarize_all_v3_supabase_only.py

➤ summarizer.py

YouTube Video Summarizer

This script provides functionality to generate Thai-language summaries of YouTube videos using OpenAI's GPT-3.5-turbo model. It processes video metadata and creates concise, news-style summaries suitable for reporting.

Features:

- Thai-language summary generation
- News-style reporting format
- Error handling and graceful degradation
- Configurable OpenAI API settings
- Legal compliance (uses only public metadata)
- Robust retry logic with exponential backoff
- Rate limit handling for API stability (โมดูล Python สำหรับ summarizer)

ฟังก์ชันหลัก: get_summarizer, summarize_video_info, summarize_thai_video, summarize_english_video, main

การใช้งาน: [CLI script](#)

ขนาดไฟล์: 26.6 KB | เส้นทาง: summarizer.py

test_ai_image_generation.py

ไฟล์ทดสอบสำหรับ ai_image_generation

ฟังก์ชันหลัก: test_supabase_connection, test_ai_image_generator, test_supabase_storage, test_supabase_ai_generator, test_ai_images_table...

การใช้งาน: [CLI script](#), [Test suite](#), [Database operations](#)

ขนาดไฟล์: 6.5 KB | เส้นทาง: test_ai_image_generation.py

test_ai_only_images.py

ไฟล์ทดสอบสำหรับ ai_only_images

ฟังก์ชันหลัก: test_backend_pipeline, test_frontend_api, test_frontend_code, test_database_views, main

การใช้งาน: [CLI script](#), [Test suite](#), [Database operations](#)

ขนาดไฟล์: 9.4 KB | เส้นทาง: test_ai_only_images.py

test_categories_pipeline.py

ไฟล์ทดสอบสำหรับ categories_pipeline

ฟังก์ชันหลัก: load_sample_data, classify_with_legacy, test_golden_set, analyze_unknown_rate, print_category_distribution...

การใช้งาน: [Test suite](#)

ขนาดไฟล์: 8.8 KB | เส้นทาง: test_categories_pipeline.py

test_compat_scores.py

ไฟล์ทดสอบสำหรับ compat_scores

ฟังก์ชันหลัก: load_sample_data, score_with_legacy, score_with_v2, extract_score, compare_scores...

การใช้งาน: [Test suite](#)

ขนาดไฟล์: 7.2 KB | เส้นทาง: test_compat_scores.py

test_idempotency_cases.py

ไฟล์ทดสอบสำหรับ idempotency_cases

ฟังก์ชันหลัก: create_mock_video_data, create_mock_image, test_case_1_incomplete_third_image, test_case_2_images_disappear, test_frontend_json_format...

การใช้งาน: [CLI script](#), [Test suite](#)

ขนาดไฟล์: 20.4 KB | เส้นทาง: test_idempotency_cases.py

test_image_fallback_fix.py

ไฟล์ทดสอบสำหรับ image_fallback_fix

ฟังก์ชันหลัก: test_frontend_data, main

การใช้งาน: CLI script, Test suite

ขนาดไฟล์: 4.3 KB | เส้นทาง: test_image_fallback_fix.py

test_image_prompt_presence.py

ไฟล์ทดสอบสำหรับ image_prompt_presence

ฟังก์ชันหลัก: load_sample_data, validate_image_prompts, print_results, main

การใช้งาน: Test suite

ขนาดไฟล์: 5.7 KB | เส้นทาง: test_image_prompt_presence.py

test_json_contract.py

ไฟล์ทดสอบสำหรับ json_contract

ฟังก์ชันหลัก: load_json_data, validate_json_contract, print_results, main

การใช้งาน: Test suite

ขนาดไฟล์: 5.9 KB | เส้นทาง: test_json_contract.py

test_no_mock_output.py

ไฟล์ทดสอบสำหรับ no_mock_output

ฟังก์ชันหลัก: scan_json_file, scan_item_for_mock_content, validate_summary_fields, check_required_fields, print_violations...

การใช้งาน: Test suite

ขนาดไฟล์: 10.5 KB | เส้นทาง: test_no_mock_output.py

test_pipeline_diagnostics.py

การทดสอบและตรวจสอบสุขภาพของ data pipeline

ฟังก์ชันหลัก: main

การใช้งาน: Test suite, Database operations

ขนาดไฟล์: 14.9 KB | เส้นทาง: test_pipeline_diagnostics.py

tests\pipeline\test_enrichment_persist.py

ไฟล์ทดสอบสำหรับ tests\pipeline\enrichment_persist

การใช้งาน: Test suite, Database operations

ขนาดไฟล์: 6.1 KB | เส้นทาง: tests\pipeline\test_enrichment_persist.py

tests\test_bnk48_enrichment.py

ไฟล์ทดสอบสำหรับ tests\bnk48_enrichment

การใช้งาน: Test suite

ขนาดไฟล์: 5.1 KB | เส้นทาง: tests\test_bnk48_enrichment.py

tests\test_build_score_details.py

ไฟล์ทดสอบสำหรับ tests\build_score_details

ฟังก์ชันหลัก: test_build_score_details

การใช้งาน: [Test suite](#)

ขนาดไฟล์: 3.7 KB | เส้นทาง: tests\test_build_score_details.py

tests\test_crypto_security.py

ไฟล์ทดสอบสำหรับ tests\crypto_security

การใช้งาน: [CLI script](#), [Test suite](#)

ขนาดไฟล์: 8.5 KB | เส้นทาง: tests\test_crypto_security.py

tests\test_images_pipeline_backend.py

ไฟล์ทดสอบสำหรับ tests\images_pipeline_backend

การใช้งาน: [Test suite](#), [Database operations](#)

ขนาดไฟล์: 7.9 KB | เส้นทาง: tests\test_images_pipeline_backend.py

tests\test_image_url_protection.py

ไฟล์ทดสอบสำหรับ tests\image_url_protection

การใช้งาน: [Test suite](#), [Database operations](#)

ขนาดไฟล์: 9.8 KB | เส้นทาง: tests\test_image_url_protection.py

tests\test_keyword_extraction.py

ไฟล์ทดสอบสำหรับ tests\keyword_extraction

ฟังก์ชันหลัก: test_basic_extraction, test_hashtag_extraction, test_multilingual_keywords,
test_featuring_artist_extraction, test_stopwords_filtering...

การใช้งาน: [Test suite](#)

ขนาดไฟล์: 8.7 KB | เส้นทาง: tests\test_keyword_extraction.py

tests\test_multilingual_enrichment.py

ไฟล์ทดสอบสำหรับ tests\multilingual_enrichment

ฟังก์ชันหลัก: test_keyword_extraction, test_platform_detection, test_bnk48_enrichment

การใช้งาน: [CLI script](#), [Test suite](#)

ขนาดไฟล์: 5.4 KB | เส้นทาง: tests\test_multilingual_enrichment.py

tests\test_score_details_deterministic.py

ไฟล์ทดสอบสำหรับ tests\score_details_deterministic

ฟังก์ชันหลัก: test_build_score_details_deterministic

การใช้งาน: [Test suite](#)

ขนาดไฟล์: 5.5 KB | เส้นทาง: tests\test_score_details_deterministic.py

test_summaries_presence.py

ไฟล์ทดสอบสำหรับ summaries_presence

ฟังก์ชันหลัก: load_sample_data, validate_summaries, print_results, main

การใช้งาน: Test suite

ขนาดไฟล์: 7 KB | เส้นทาง: test_summaries_presence.py

update_video_data.py

TrendSiam Video Data Updater

=====

Safely updates video view counts and categories from YouTube Data API while maintaining full compatibility with existing TrendSiam system and security policies.

Features:

- Fetches latest view counts, likes, comments from YouTube API
- Re-classifies categories using existing TrendSiam logic
- Defensive programming to prevent data loss
- Backup and rollback functionality
- Rate limiting and API quota management
- Full validation and error handling

Requirements:

- YOUTUBE_API_KEY in .env file
- Existing thailand_trending_summary.json file
- All TrendSiam dependencies installed

Usage:

python update_video_data.py [--dry-run] [--backup-only] [--restore]

Security:

- Validates all API responses
- Creates automatic backups before changes
- Maintains data integrity with extensive validation
- Respects YouTube API terms of service
- Implements rate limiting and error recovery (โมดูล Python สำหรับ update_video_data)

ฟังก์ชันหลัก: main

การใช้งาน: CLI script

ขนาดไฟล์: 23.3 KB | เส้นทาง: update_video_data.py

utils__init__.py

Utility functions for the TrendSiam pipeline (โมดูล Python สำหรับ utils__init__)

ขนาดไฟล์: 287 B | เส้นทาง: utils__init__.py

utils\keyword_extractor.py

Advanced keyword extraction for multilingual content (โมดูล Python สำหรับ utils\keyword_extractor)

ขนาดไฟล์: 14 KB | เส้นทาง: utils\keyword_extractor.py

utils\safe.py

Safe utilities for protecting existing data during pipeline operations (โมดูล Python สำหรับ utils\safe)

ฟังก์ชันหลัก: truthy_url, set_if_truthy, get_safe_url

ขนาดไฟล์: 1012 B | เส้นทาง: utils\safe.py

⇒validate_e2e_fixes.py

Validate E2E fixes are working correctly (โหลด Python สำหรับ validate_e2e_fixes)

การใช้งาน: Database operations

ขนาดไฟล์: 3.5 KB | เส้นทาง: validate_e2e_fixes.py

⇒validate_fixes.py

[DEPRECATED] Use validate_e2e_fixes.py instead

Validate that the pipeline 'DB 'API 'Frontend flow is working correctly (โหลด Python สำหรับ validate_fixes)

การใช้งาน: Database operations

ขนาดไฟล์: 4 KB | เส้นทาง: validate_fixes.py

⇒youtube_api_fetcher.py

ดึงข้อมูลจาก YouTube Data API

ฟังก์ชันหลัก: main

การใช้งาน: CLI script

ขนาดไฟล์: 11.6 KB | เส้นทาง: youtube_api_fetcher.py

⇒youtube_fetcher.py

การดึงข้อมูลและ web scraping สำหรับ youtube_fetcher

ฟังก์ชันหลัก: main

การใช้งาน: CLI script

ขนาดไฟล์: 17.6 KB | เส้นทาง: youtube_fetcher.py

⇒Frontend/src/app

⇒Frontend/src/app/weekly-report

⇒Frontend/src/lib/weekly

⇒Frontend/src/app/api/weekly/pdf

ก.2 สคริปต์สำคัญ

สคริปต์หลักสำหรับการดึงข้อมูล

- summarize_all_v2.py - สคริปต์หลักสำหรับดึงข้อมูลจาก YouTube
- youtube_api_fetcher.py - ตัวดึงข้อมูล API เฉพาะ
- ai_image_generator_v2.py - สร้างภาพประกอบด้วย AI

สคริปต์ทดสอบและตรวจสอบ:

- test_pipeline_diagnostics.py - ตรวจสอบสุขภาพ pipeline
- acceptance_tests.py - ทดสอบการยอมรับ
- security_audit.py - ตรวจสอบความปลอดภัย

สคริปต์ตั้งค่าและบำรุงรักษา

- setup_environment.py - ตั้งค่าสภาพแวดล้อม
- clean_trending_data.py - ทำความสะอาดข้อมูล
- popularity_scorer.py - คำนวณคะแนนความนิยม

ภาคผนวก ข: รายการ API และเส้นทาง

ข.1 API Endpoints

GET	/api/health	ตรวจสอบสถานะระบบโดยรวม	การยืนยันตัวตน: ไม่ต้อง	ผลลัพธ์: JSON status object
GET	/api/health/db	ตรวจสอบการเชื่อมต่อฐานข้อมูล	การยืนยันตัวตน: ไม่ต้อง	ผลลัพธ์: Database connection status
GET	/api/weekly/diagnostics	ข้อมูลการวินิจฉัยระบบรายงานรายสัปดาห์	การยืนยันตัวตน: ไม่ต้อง	ผลลัพธ์: Diagnostic information
GET	/api/weekly/pdf	สร้างและดาวน์โหลดรายงาน PDF รายสัปดาห์	การยืนยันตัวตน: ไม่ต้อง	ผลลัพธ์: PDF file
GET	/api/weekly/data	ข้อมูลรายงานรายสัปดาห์ในรูปแบบ JSON	การยืนยันตัวตน: ไม่ต้อง	ผลลัพธ์: Weekly snapshot data
POST	/api/admin/revalidate	บังคับให้ระบบโหลดข้อมูลใหม่	การยืนยันตัวตน: x-revalidate-secret header	ผลลัพธ์: Revalidation result
GET	/api/env-check	ตรวจสอบการตั้งค่า environment variables	การยืนยันตัวตน: ไม่ต้อง	ผลลัพธ์: Environment status

ข.2 หน้าเว็บไซต์

/	Public	หน้าแรก - แสดงข้อมูลข่าวแนวโน้มล่าสุด
/weekly-report	Public	

/dev-dashboard

Development

แดชบอร์ดสำหรับนักพัฒนา

/legal

Public

ข้อมูลทางกฎหมาย

/privacy

Public

นโยบายความเป็นส่วนตัว

/terms

Public

ข้อกำหนดการใช้งาน

ข.3 คำสั่ง CLI สำคัญ

```
# Frontend Commands
npm run dev                # development server
npm run build              # production build
npm run snapshot:build:publish # build & publish snapshot
npm run snapshot:test:selection # test selection

# Backend Commands
python summarize_all_v2.py --20 # summarize last 20 days
python test_pipeline_diagnostics.py # test pipeline
python security_audit.py # security audit

# Health Checks
curl http://localhost:3000/api/health
curl http://localhost:3000/api/weekly/diagnostics
```

ข.4 Environment Variables สำคัญ

ตัวแปรสภาพแวดล้อมที่จำเป็นสำหรับการทำงานของระบบ

```
NEXT_PUBLIC_SUPABASE_URL      # URL - Supabase project
NEXT_PUBLIC_SUPABASE_ANON_KEY # Anon key *30#1
SUPABASE_SERVICE_ROLE_KEY     # Service role *30#1
OPENAI_API_KEY                 # API key *30#1
YOUTUBE_API_KEY                # API key *30#1
REVALIDATE_SECRET              # Secret *30#1
```