GenAI Workflow：
打造智能化技術趨勢洞察系統

2025.08.09 張維峻

Today’s agenda

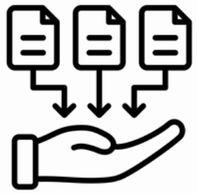
Why Collect New Technologies?

The Rapid Development of Modern Technology and Generative AI

Why a Technology Trend Insight System is Needed

The Value of an Insight System

System Pain Points



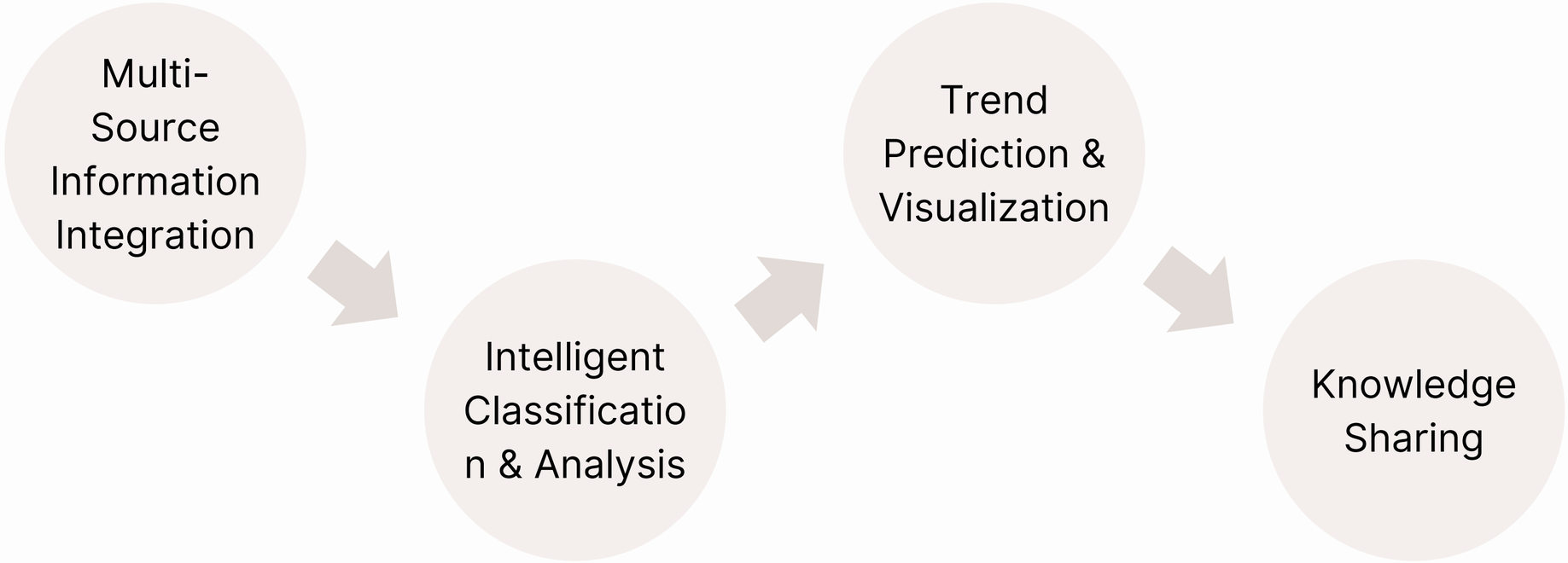
Data Collection

Data Storage

Data Analysis

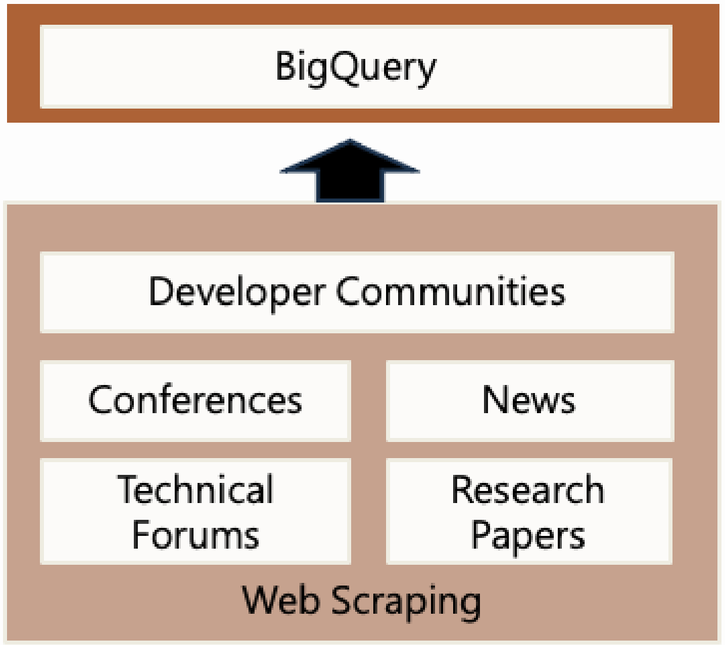
Application Layer

Architecture Blueprint



(A) Open-Source Data Pipeline Construction

"Efficiently collect technical papers, industry reports, and conference information."



Open-Source Crawler Frameworks

- Demonstrate how to use tools like Selenium and Crawl4ai for efficient data collection.

Data Acquisition Strategy

- Identify key technical domains and establish diverse information sources.

Automated Monitoring

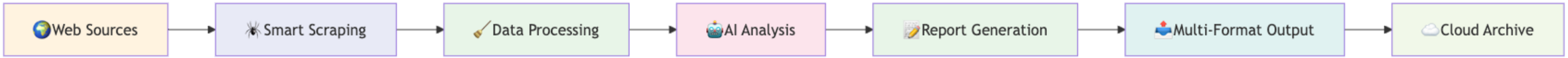
- Use RSS subscriptions, API integrations, etc., to get the latest tech trends in real time.

Key Takeaway

- Data acquisition strategy and Schema design considerations.

(B) GenAI Intelligent Analysis Workflow

"Designing & Deploying Scalable GenAI Workflows"



🤝 Advanced Model Integration

- Large Language Models (LLMs): Leveraging state-of-the-art models like GPT-4o and Gemini 2.5 Pro for novel content generation.
- Multimodal Capabilities: Integrating support for diverse data types, including text, images, and speech.

⚙️ End-to-End Automated Process

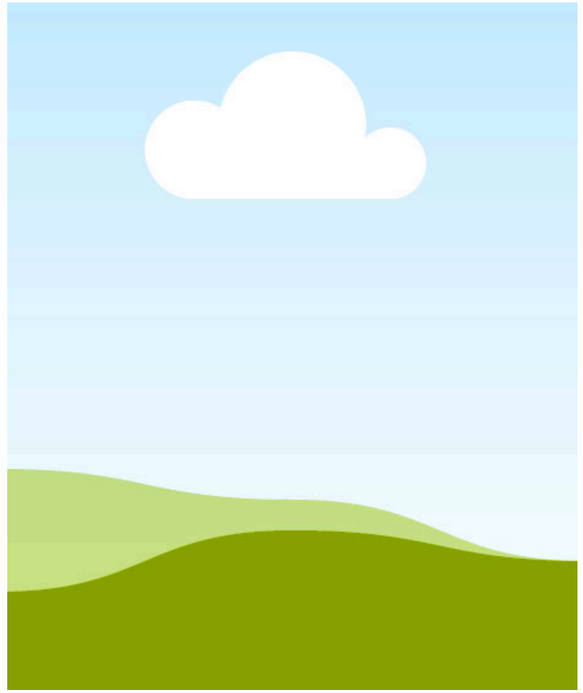
1. Data Cleaning & Preprocessing 🧹
2. Automated Feature Extraction 🔍
3. AI-Powered Model Inference 🤖
4. Generation of Structured Technical Insight Reports 📊

🛠️ Critical Optimization Techniques

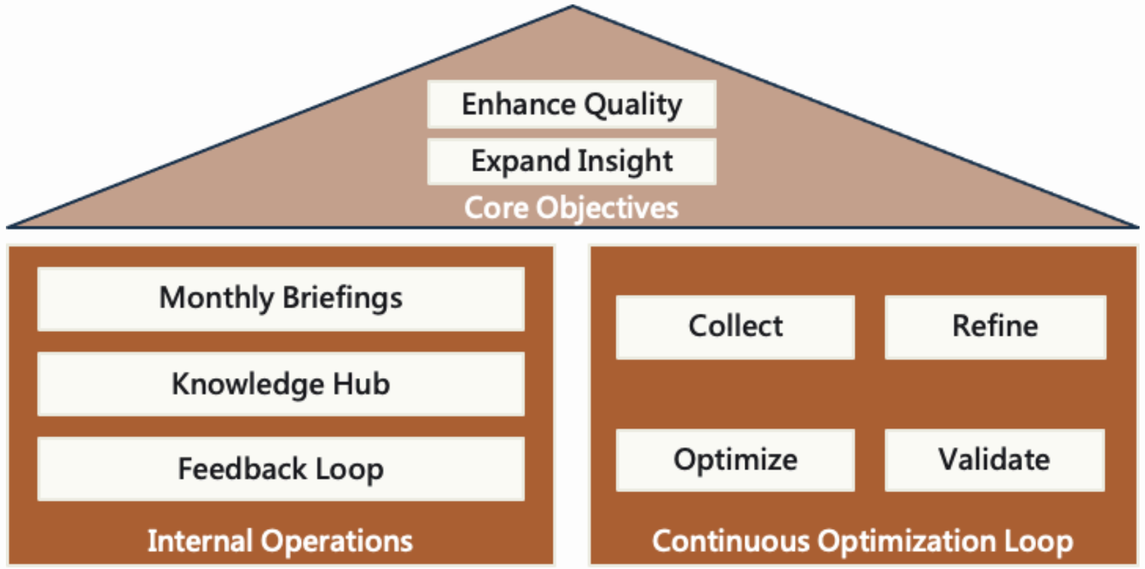
- Prompt Engineering ✏️: Methodically crafting prompts to refine and enhance the quality of model outputs.

(C) Interactive Trend Visualization

Section	Details
🛠️ Technical Delivery	We generate single-file, portable HTML reports with fully embedded interactive features, ensuring they work anywhere without dependencies.
📊 Core Functionality	• **Trend Evolution Tracking:** Dynamically visualize the rise and fall of technological popularity over time. • **Multi-Dimensional Semantic Visualization:** Map the underlying connections and application relationships between different technologies.
💡 Project Value	• Simplifies Complexity: Provides an at-a-glance understanding of intricate market and technology trends. • Enhances Collaboration: Reports can be seamlessly embedded in websites and internal systems or shared directly to foster team alignment.



(D) Open-Source Community Collaboration and Iteration



Core

- Improve system and reports through open collaboration and continuous iteration.
- Broaden knowledge coverage to strengthen decision support.

Internal Operations

- Share the latest tech trend reports in department meetings.
- Provide reports as a go-to reference for tech and industry insights.
- Collect team feedback on report accuracy and presentation.

Continuous Optimization Loop

Takeaways

Build a fully automated tech trend analysis system based on an open-source tech stack.

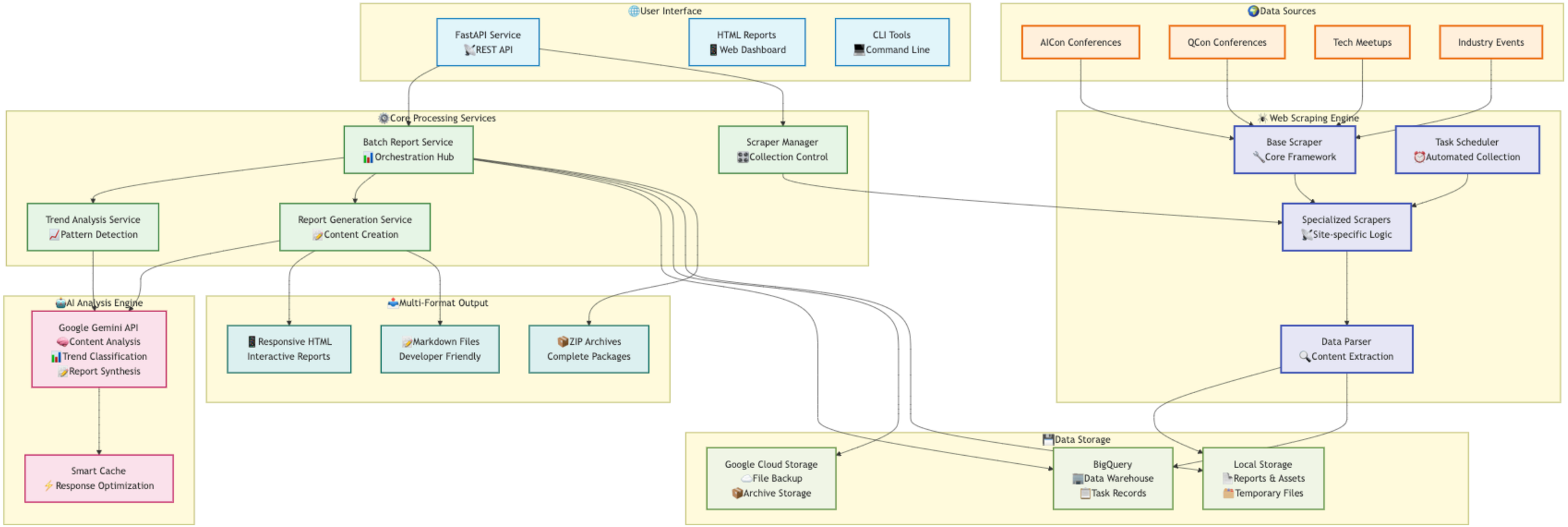
Open-Source Data Pipeline: Collection via Selenium / Crawl4ai, storage, and retrieval with BigQuery.

GenAI Intelligent Analysis: Automated end-to-end workflow using GPT-4o / Gemini 2.5 Pro.

Interactive Visualization: HTML interactive reports supporting multi-dimensional filtering and in-depth analysis.

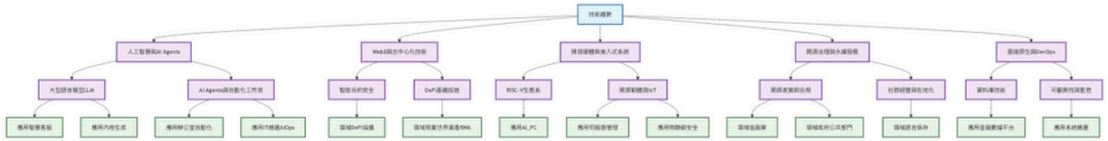
Community Collaboration & Iteration: Feedback and co-creation to accelerate innovation and wider application.

Architecture diagram



Architecture diagram

Coscup 2025 趨勢總覽



趨勢一：人工智慧與 AI Agents 的深化與普及

子趨勢：大型語言模型（LLM）與生成式 AI

- **總覽與建議**：
  - 趨勢從通用模型轉向專用化與效率化。透過模型微調（fine-tuning）、極致量化（如 BitNet）和特定領域資料（如國會逐字稿、台灣在地語料），打造更精準、更輕量的 AI 應用。開發者應關注如何將開源模型（如 Gemma）與自有資料結合，解決特定場景問題。
- **應用場景**：
  - **智慧客服與助理**：結合 n8n、LangChain 等工具打造自動化問答系統。
  - **內容生成與分析**：自動生成國會逐字稿、分析網路監控數據、撰寫部落格。
  - **企業內部應用**：處理個資過濾、文件摘要等敏感資訊。
  - **開發輔助**：利用 Vibe Coding、Copilot 等工具加速程式碼生成與原型開發。
- **潛在影響**：
  - 大幅降低特定領域 AI 應用的開發門檻，實現「人人皆可 AI」。
  - 推動從雲端 AI 到裝置端（Edge AI）的轉移，保護資料隱私。
- **關鍵挑戰**：
  - **模型幻覺（Hallucination）**：生成的內容仍需驗證，需要建立可控的工作流。
  - **運算資源與成本**：模型量化與優化是普及化的關鍵。
  - **資料隱私與安全**：企業導入時需謹慎處理敏感資料。

Thanks!