**系統架構雛型**

統計資訊網係以收集由各觀光相關單位回報之資料彙集後以特定條件查詢產生數據後呈現給民眾或研究單位，或是將原始資料產生給研究單位使用，因此資料的輸入、處理與輸出將是本系統的技術重點工作，其系統架構雛型如下圖所示。



系統架構分為資料管理平台 (Data Management Platform)、報告平台 (Reporting Platform) 兩大部份，並以奇豐應用服務平台 (Application Services Platform) 所打造，平台間並共用安全性、帳戶管理、日誌等共用模組。

|  |  |
| --- | --- |
| 模組 | 說明 |
| 資料管理平台 | 負責執行觀光統計原始資料之匯入、檢核、處理與輸出等工作。 |
| 報告平台 | 由資料庫查詢特定之資料集或數據後，呈現於不同的格式，包含原始資料的輸出或資料視覺化之圖表等。 |
| 共用模組 | 負責系統的共用功能，如安全機制、帳戶管理、日誌記錄等功能。 |

**資料分析**

為進行統計資料庫之建構，奇豐將於專案初期對現有資料庫進行分析，並輔以需求訪談確認單位之統計資料需求後，設計本專案之統計資料庫，並確認資料來源類型，用以實作資料來源連結器 (Data Adapter)、相關統計數據所需之資料特徵 (Data Attributes)，以及實作資料流處理中之必要資料處理工作 (Tasks)。

資料庫的設計將視資料量與分析之方法，調整為以資料倉儲或大數據為基礎之設計架構。

**數據建立管理**

資料管理平台負責資料的匯入，包含原始資料的匯入管道，如API、原始檔案 (CSV、Excel或文字檔等) 或Access資料庫等來源，每個資料來源都有一個資料來源連結器 (Data Adapter) 負責進行資料輸入端的載入與檢驗，再進入資料匯入的ETL (Extract, Transform and Load) 階段，針對目標資料庫或資料倉儲的結構進行轉換、資料檢驗與輸入等工作，此工作為批次的管線化處理 (Batch Pipeline-based Processing)。



在資料匯入管線內採用資料流分析機制 (Streaming Analytics) 技術，定義資料匯入時需萃取的資料特徵 (Data Attribute)，以提高資料匯入的彈性，減少資料匯入時重覆執行的負擔。



資料匯入於資料庫內後，除可供數位儀表板之統計數據查詢依據外，亦可依需求進行匯出，匯出的格式以現行政府開放資料格式為準，包含CSV、XML、JSON、TXT等格式或Excel格式等，供學術研究單位參考使用，資料內容可以是原始資料或已經處理過後之資料。

除供研究機構使用之資料外，針對一般民眾，系統亦提供報表之下載，以資訊圖表 (Chart) 或資料視覺化 (Data Visualization) 方式呈現，讓民眾得以快速了解報表之數據內容與其意義。



另外，資料匯入時所需之安全控管機制，由共用模組內之單一簽入系統 (Single Sign On) 提供，使用者只需記憶一組帳戶即可用於各類資料之匯入工作。

**數據檢核機制**

數據檢核機制 (Data Validation Tasks) 於資料來源載入與資料管線作業中執行，目的在於檢查資料是否有出現無效值、空值或非預期範圍的情況，如檢核機制發現資料出現異常狀態時，將會使用預設值、回報錯誤狀況、保留原始數值等策略進行回應，原則上必要的欄位會以回報錯誤狀況為主，次要欄位則是以預設值為主，保留原始值或回報錯誤狀況為輔。

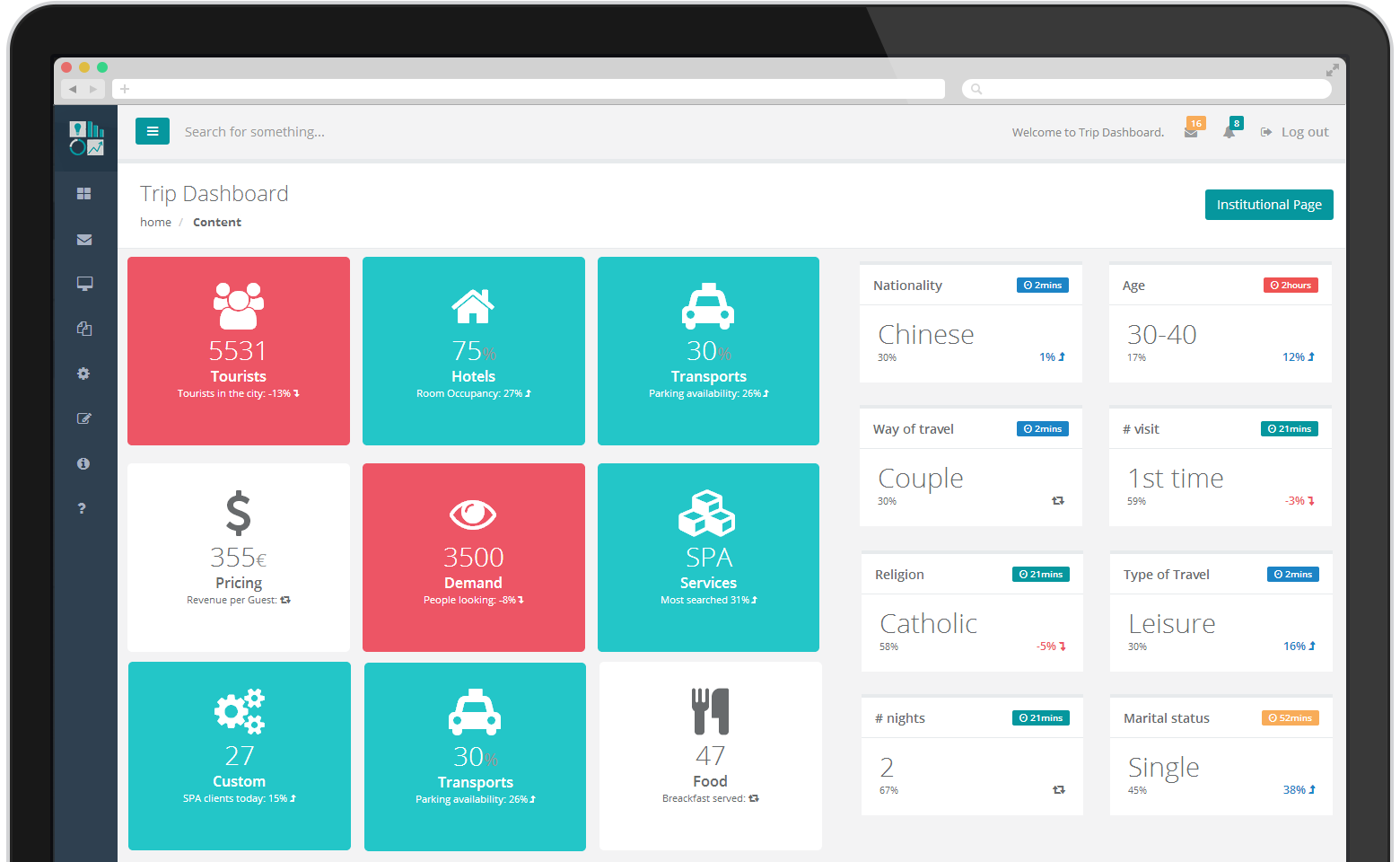
資料檢驗之異常狀況與動作處理均會於日誌中留下記錄以供備查之用。

**數位儀表板**

數位儀表板 (Dashboard) 為報表平台功能之一。

統計資料庫內之資料可選擇特定之主題查詢 (Subject-based Queries)，由前台以儀表板之模式產生，並可以比較區間之方式呈現，例如「來臺觀光之國家人口的成長比例」、「臺灣各機場出國之人數比例與年度比較」等特定指標，直接於畫面呈現，讓民眾可以在最短時間查詢到相關資料，儀表板之查詢條件與呈現方式可於系統後台指定。同時針對無法直接呈現予民眾之業務統計類型資料，亦可由系統後台之數位儀表板呈現提供。

隨著觀光統計資料類型的增加，數位儀表板可掌握的資料與統計數據類型也會逐步增加。

[](https://www.google.com/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=2ahUKEwiNysO7lNnZAhVHqVQKHUi3C5sQjRx6BAgAEAU&url=http://tripdashboard.awr.pt/&psig=AOvVaw2Tt090hJQlUqZVpkBP-qvj&ust=1520475898002163)

針對數位儀表板之數據查詢，於系統中採用快取機制，當有新資料匯入或是需重新查詢數據時才會執行，並存入快取區中，日後再次的查詢將直接由快取取得，降低資料庫的查詢負載。

**未來資料應用**

1. 結合智慧裝置數據

資料管理平台內的API連接器 (API Data Adapter) 可依照不同的資料來源建立自動收集資料的API端點 (Endpoint)，系統管理人員可依照需求建立不同的API端點交由各系統或IoT設備連結 (例如回傳即時遊客人數)，定時回報資料內容，並依照需要之頻率設定回傳之週期，除可逐步免除手動上傳之程序，也讓資料更新速度加快且自動化。

未來若邊界運算 (Edge Compute) 成熟時，可進一步由IoT設備或外部系統取得更多細部資料後自動回傳，讓統計資料庫內的資料更加多元化。

1. 大數據分析

統計資料庫的量達到一定程度後，分析方法就需要改變，改採用適合大數據的資料分析方法執行，以負擔來自各個不同系統或設備所傳回的資料，包含採用分散式資料分析架構、資料流分析等大數據分析常用作法，使資料的分析能貼近即時。

1. 模式建立與回饋

大數據分析模式建立後，即可對資料進行模式分析，包含資料探勘或統計模型之分析，結合機器學習 (Machine Learning) 的技術，發展並訓練觀光相關之資料分析模式，並回饋模式分析結果予各觀光單位作為發展參考，或是進一步導入人工智慧 (AI)，發展如觀光機器人 (Tour Bot) 給予旅遊建議等自動化工具，以提升觀光服務品質。