

海軍大氣海洋局 114-116 年度委託研究計畫構想書

計畫名稱：異質大數據資料處理模組發展案之研究		計畫期限：114-116 年
提案單位：海軍大氣海洋局 聯絡人：吳旻修上尉 電話：0986181715		
項次	項目	內容
1	計畫目的	透過結合數據科學及人工智慧技術，開發自動化資料處理模組，提升本軍在異質與巨量資料環境下的資料處理能力與速度。透過多樣化資料來源(包含文字情資、圖像、音響及地理空間等)，建立一套高效、準確的資料清洗、標籤及儲存流程，以利作戰演訓、戰情分析和情報判斷等軍事應用。最終，目標為建構一個先進的異質大數據資料庫系統，能夠高效管理和利用各類資料，並提供靈活地檢索和分析功能，支持本軍的業務需求和未來發展。
2	研究議題	<p>1. 總計畫暨子計畫一：異質大數據資料處理方法研析</p> <p>(1)第一年：針對大氣海洋局目前的資料類型及格式，如環境資訊(氣象、水文、海圖、磁力等)、影像資訊(衛星照片、光學影像)、文字資訊、電磁資訊(發射及接收頻率、AIS、ADB-S)、音頻資訊(量測及理論音頻)，進行研析並尋求合宜處理方法。</p> <p>(2)第二年：透過資料科學的資料邏輯處理流程，進行數據前處理(清理和格式化數據)、特徵提取(提取能夠代表數據特性的特徵向量)、模型訓練(使用適當的機器學習或深度學習模型進行訓練)、分類及分群(根據特徵向量進行分類或分群)、離型模型評估及調優(評估模型的效果，並進行調參和優化)等方法，對異質資料處理進行研究分析及評估，以建立異質大數據的自動化處理方法。</p> <p>(3)第三年：同時評估後續產出資料庫系統的形式，以利巨量資料完成處理分析後得以運用適當資料庫系統進行管理、存取運用。</p> <p>2. 子計畫二：資料自動化處理程序</p> <p>(1)第一年：結合異質處理方法研析，在這些異質類型資料於開始接收時，執行合宜、正確、有效的處理、清洗、標籤及儲存。</p> <p>(2)第二年：規劃從資料接收開始、資料清洗、資料標籤、資料轉譯、資料儲存等處理程序，運用人工智能技術，</p>

		<p>透過適當工具及程式語言編寫處理流程。</p> <p>(3)第三年：結合大海局需求，參考現有轉譯文本，建立需轉譯文本資料之表單格式，並自動產出，資料型態包含：txt, netCDF, xyz, Geotiff, jpg, pdf, shp, kml, dxf, meta, png,mp4, mov, wmv, word 檔, execl 檔, json, wav, mp3, csv 等。</p> <p>3.子計畫三：大數據異質資料庫系統建立研究</p> <p>(1)第一年：由於本計畫所處理的資料為異質性資料，包含圖片，音訊，影象，文字...等。傳統關聯式資料庫不足以應付這樣的資料樣態。我們將評估及規劃異質大數據處理處理模組所需的合宜資料庫與資料庫管理系統(例如 noSQL 資料庫)，並建立先導試驗性質之微型雛形系統。建立各異質資料原始數據及分析後續資料庫並可與現有資料庫系統鏈結。針對各異質資料建立對應的 meta data，以及相關的索引資訊，並評估適合的搜尋引擎技術(如 Elasticsearch)，以支援異質性資料庫的搜尋與存取。完成數據自動化處理的海洋環境資料庫，包含環境資訊(氣象、水文、海圖、磁力等)、影像資訊(衛星照片、光學影像)、文字資訊、電磁資訊(發射及接收頻率、AIS、ADB-S)、音頻資訊(量測及理論音頻)等資料可供程式檢索，可根據使用者輸入的關鍵字或特定參數，檢索所需的資料，並輸出成轉譯文本，以利大氣海洋局業管人員進行資料運用所需。</p> <p>(2)第二年：設計使用者介面，讓使用者可以方便的存取並搜尋異質性資料。使用者介面前期將以 web-based 的介面設計，待應用成熟後，逐步提供多平台(web、手機、平板)的自適應前端架構。</p> <p>(3)第三年：制定全面的測試計劃，包括單元測試、整合測試、性能測試和安全測試，確保系統的穩定性和可靠性。</p>
3	運用構想	<p>1、案內建置之資料庫可另提供於水下物聯網及預測式人工智慧情資判斷系統運用，俾利提供本軍後續聯合情監偵系統建立。</p> <p>2、模組建置將考量後續運用靈活度，未來如有新格式資料匯入，將可擴充此模組，以利資料存取。</p>

		3、結案前需解繳相應儲存設備、使用者介面開發及使用所需硬體設備及智慧紙本情資掃描轉譯系統，供提案單位執行測試、驗證、展示及儲放研究成果。
4	成果運用	<p>1、提供後續聯合情監偵系統建置所需異質料處理模組，清洗並統一資料格式，以減少整體資料處理程序人員介入比例。</p> <p>2、提供紙本情資數位化程序，減少人工鍵入所需時間，以提升整體作業效能。</p>