

資料工程人員職能基準

版本	職能基準代碼	職能基準名稱	狀態	更新說明	發展更新日期
V3	SMS2529-001v3	資料工程人員	最新版本	略	2024/12/15
V2	SMS2529-001v2	資料工程人員	歷史版本	已被《SMS2529-001v3》取代	2021/12/31
V1	SMS2529-001v1	資料工程人員	歷史版本	已被《SMS2529-001v2》取代	2019/12/19

職能基準代碼		SMS2529-001v3			
職能基準名稱 (擇一填寫)		職類			
		職業	資料工程人員		
所屬 類別	職類別	科學、技術、工程、數學 / 數學及科學		職類別代碼	SMS
	職業別	其他資料庫及網路專業人員		職業別代碼	2529
	行業別	專業、科學及技術服務業 / 研究發展服務業		行業別代碼	M7230
工作描述		依需求確認資料蒐集內容、建立資料介接流程並進行資料處理、轉置，以及數據平台之維運管理。			
基準級別		4			

主要職責	工作任務	工作產出	行為指標	職能 級別	職能內涵 (K=knowledge 知識)	職能內涵 (S=skills 能力)
T1 確認任務需求及資料特性	T1.1 確認需求	O1.1.1 需求確認書	P1.1.1 與需求單位面談，以決定資料分析或處理之目的或需求 P1.1.2 將文件、紀錄等相關之需求資訊，與需求單位確認。	3	K01 資料應用知識 K02 組織資料蒐集處理相關政策	S01 資料蒐集能力 S02 資料分析能力 S03 溝通能力
	T1.2 建立資料蒐集清單	O1.2.1 資料清冊 (含資料)	P1.2.1 以文件化方式記錄各項欲蒐集資料之範圍、項目、數量、資料原始型態、資料提供之形式及其他相關必要資訊，並產出資	3	K01 資料應用知識 K02 組織資料蒐集處理相關政策	S01 資料蒐集能力 S02 資料分析能力 S03 溝通能力

主要職責	工作任務	工作產出	行為指標	職能級別	職能內涵 (K=knowledge 知識)	職能內涵 (S=skills 能力)
		工作流 程)	料清冊。			
T2設計與 建置數據 管理機制	T2.1確認 儲存方案	O2.1.1技 術設計文 件	P2.1.1依據資料量與特性，評估各儲存方案之資料接收及讀取方式、效率、可儲存容量以及擴充機制等所需規格資訊。 P2.1.2依據上述資料進行概念驗證 (POC)，並記錄驗證過程中之相關數據，比對各儲存方案之規格資訊。 P2.1.3依據需求，建議最適之資料儲存方案，將技術規格設計細節製作為文件，並取得需求單位之同意。	4	K03程式語言 K04資料庫 K05資料儲存機制 K06資訊網路概論 K07伺服器架構	S04資料庫技術應用 S05程式語言應用 S06分散運算與儲存平台技術應用 S07需求整合能力 S08資料儲存分析平台操作 S09分散運算與儲存技術應用 S10資料處理、串流分析與技術應用
	T2.2進行 數據平台 建置	O2.2.1數 據平台安 裝文件 O2.2.2數 據平台權 限設定文 件	P2.2.1確認數據平台安裝之所需軟硬體清單。 P2.2.2製作安裝程序與配置之相關文件 (如安裝 SOP)。 P2.2.3建置數據平台，安裝軟體、配置數據平台環境組態與服務。 P2.2.4驗證數據平台軟體與服務功能，確保可正確運作。 P2.2.5依最小權限原則，設定數據平台權限。	4	K03程式語言 K04資料庫 K05資料儲存機制 K06資訊網路概論 K07伺服器架構 K08資訊安全概論	S04資料庫技術應用 S05程式語言應用 S06分散運算與儲存平台技術應用 S08資料儲存分析平台操作 S09分散運算與儲存技術應用 S10資料處理、串流分析與技術應用 S11作業系統操作能力
	T2.3進行 資料存放	O2.3.1資 料配置文 件	P2.3.1依據欲儲存之資料類型、來源或需求，設計資料集或資料標籤。 P2.3.2設定各資料集組態，例如資料存放位置、	4	K03程式語言 K04資料庫 K05資料儲存機制	S04資料庫技術應用 S05程式語言應用 S06分散運算與儲存平台技術應

主要職責	工作任務	工作產出	行為指標	職能級別	職能內涵 (K=knowledge 知識)	職能內涵 (S=skills 能力)
			資料欄位名稱、資料長度、資料最大使用空間等必要資訊。 P2.3.3進行資料匯入作業，並驗證資料存放位置與組態設定一致性。		K06資訊網路概論 K07伺服器架構	用 S08資料儲存分析平台操作 S09分散運算與儲存技術應用 S10資料處理、串流分析與技術應用
	T2.4建立介接機制	O2.4.1介接機制文件	P2.4.1確認資料來源與型式與資料接收後之儲存方式，包含儲存位置，切分方式，資料格式與所需內容。 P2.4.2依資料需求，評估建立資料介接技術進行資料蒐集或內容過濾，並建立資料儲存機制或自動化技術，存放至數據平台。 P2.4.3驗證資料介接的正確性。	4	K03程式語言 K04資料庫 K05資料儲存機制 K06資訊網路概論 K07伺服器架構	S04資料庫技術應用 S05程式語言應用 S06分散運算與儲存平台技術應用 S08資料儲存分析平台操作 S09分散運算與儲存技術應用 S10資料處理、串流分析與技術應用 S12整合測試能力 S13網路及資料介接技術應用
T3數據處理與轉置	T3.1進行資料萃取	O3.1.1資料集或資料表	P3.1.1由數據平台中確定所需資訊，並確認資料分析平台之資料接取機制。 P3.1.2使用適當的資料查詢語法或開發所需程式，將所需資料接取至資料分析平台。 P3.1.3在資料分析平台上，計算資料接取成效。	4	K01資料應用知識 K03程式語言 K04資料庫 K09資料結構 K10統計學 K11資料處理 K12ETL 機制	S05程式語言應用 S10資料處理、串流分析與技術應用 S14資料分析工具應用 S15資料處理語法應用 S16ETL 相關軟體操作
	T3.2進行	O3.2.1資	P3.2.1針對數值型資料執行所需之統計分析，取	4	K01資料應用知識	S05程式語言應用

主要職責	工作任務	工作產出	行為指標	職能級別	職能內涵 (K=knowledge 知識)	職能內涵 (S=skills 能力)
	資料描述	料描述資訊	<p>得值域資訊，如資料筆數、最大值、最小值、中位數、欄位定義等必要資訊。</p> <p>P3.2.2針對非數值資訊，取得相關資訊，如統計關鍵字詞的個數、類別數、分佈狀態集中或分散等。</p>		K03程式語言 K04資料庫 K09資料結構 K10統計學 K11資料處理 K12ETL 機制	S10資料處理、串流分析與技術應用 S14資料分析工具應用 S15資料處理語法應用 S16ETL 相關軟體操作
	T3.3處理異常資料	O3.3.1調整完成之資料集或資料表	<p>P3.3.1以適當之資料查詢語法或開發所需程式，逐筆驗證資料完整性，並針對不完整之資料，以適當方式進行修正或補正，並將修正結果回存。</p> <p>P3.3.2以適當之資料查詢語法或開發所需程式，過濾出重複資料，並向相關人員確認處理方式，進行修正後將修正結果回存。</p> <p>P3.3.3以適當之資料查詢語法或開發所需程式，過濾會影響資訊分析之異常資訊 (如格式異常、數值異常、類別異常等)，進行修正後將修正結果回存。</p>	4	K01資料應用知識 K03程式語言 K04資料庫 K09資料結構 K10統計學 K11資料處理 K12ETL 機制	S05程式語言應用 S10資料處理、串流分析與技術應用 S14資料分析工具應用 S15資料處理語法應用 S16ETL 相關軟體操作
	T3.4轉置資料	O3.4.1轉置完成之資料集或資料表	<p>P3.4.1與相關人員討論，確認後續分析所需之資料轉置需求，並完成紀錄。</p> <p>P3.4.2以適當之資料查詢語法或開發所需程式，將資料內容轉換至可分析之型態，包含對資料進行分類、置換、轉置，為數值資料進行分類、貼標等必要作業。</p>	4	K01資料應用知識 K03程式語言 K04資料庫 K09資料結構 K10統計學 K11資料處理	S05程式語言應用 S10資料處理、串流分析與技術應用 S14資料分析工具應用 S15資料處理語法應用 S16ETL 相關軟體操作

主要職責	工作任務	工作產出	行為指標	職能級別	職能內涵 (K=knowledge 知識)	職能內涵 (S=skills 能力)
			<p>P3.4.3調整資料轉置作業，確保資料轉置效率與正確性等最佳化，並符合作業之需求。</p> <p>P3.4.4將轉置完成之資料，依據資料模型存入數據平台。</p>		K12ETL 機制	
T4數據平台維運	T4.1建立數據平台維運機制	<p>O4.1.1監控機制設定與組態文件</p> <p>O4.1.2監控日誌</p>	<p>P4.1.1建立數據平台的效能與容量標準。</p> <p>P4.1.2針對數據平台或資料分析平台之系統資源進行管理。</p> <p>P4.1.3建立監控機制，以確保數據平台符合效能與容量標準。</p> <p>P4.1.4將相關監控機制留存於日誌，以追蹤未達標準之情況，並主動進行改善。</p>	3	<p>K03程式語言</p> <p>K04資料庫</p> <p>K05資料儲存機制</p> <p>K08資訊安全概論</p> <p>K13個資保護相關法規</p> <p>K14伺服器管理</p> <p>K15系統日誌管理及監控</p>	<p>S05程式語言應用</p> <p>S06分散運算與儲存平台技術應用</p> <p>S11作業系統操作能力</p> <p>S17資料儲存分析平台操作</p>
	T4.2確保數據平台安全性	O4.2.1數據平台權限申請程序	<p>P4.2.1依組織資訊安全政策與規範，建立數據平台使用之權限申請程序。</p> <p>P4.2.2建立數據平台資料安全之相關規範，如必要之加密、去識別化、遮罩、資料攜出提供原則等，以確保符合現行相關法規與組織內之相關要求。</p> <p>P4.2.3建立數據平台活動監控機制，包含使用操作、運作情況與網路活動監控等，以確保數據平台安全性。</p>	3	<p>K03程式語言</p> <p>K04資料庫</p> <p>K05資料儲存機制</p> <p>K08資訊安全概論</p> <p>K13個資保護相關法規</p> <p>K14伺服器管理</p> <p>K15系統日誌管理及監控</p>	<p>S05程式語言應用</p> <p>S06分散運算與儲存平台技術應用</p> <p>S11作業系統操作能力</p> <p>S17資料儲存分析平台操作</p>
	T4.3進行數據平台的備援	O4.3.1資料備援管理文件	<p>P4.3.1建立數據平台備援程序。</p> <p>P4.3.2設定或建立數據平台備援作業。</p> <p>P4.3.3定期進行資料復原演練，並確認資料還原</p>	3	<p>K03程式語言</p> <p>K04資料庫</p> <p>K05資料儲存機制</p>	<p>S05程式語言應用</p> <p>S06分散運算與儲存平台技術應用</p>

主要職責	工作任務	工作產出	行為指標	職能級別	職能內涵 (K=knowledge 知識)	職能內涵 (S=skills 能力)
		O4.3.2資料還原演練紀錄	正確性。		K08資訊安全概論 K13個資保護相關法規 K14伺服器管理 K15系統日誌管理及監控 K16備份、復原	S11作業系統操作能力 S17資料儲存分析平台操作
	T4.4留存數據平台操作軌跡	O4.4.1數據平台日誌管理文件 O4.4.2數據平台操作日誌	P4.4.1建立數據平台日誌留存程序。 P4.4.2配置數據平台或開發所需程式，針對數據平台操作、運作之相關軌跡以日誌方式記錄，日誌內容應包含人、事、時、地、物等。	3	K03程式語言 K04資料庫 K05資料儲存機制 K08資訊安全概論 K13個資保護相關法規 K14伺服器管理 K15系統日誌管理及監控	S05程式語言應用 S06分散運算與儲存平台技術應用 S11作業系統操作能力 S17資料儲存分析平台操作

職能內涵 (A=attitude 態度)

- A01主動積極：不需他人指示或要求能自動自發做事，面臨問題立即採取行動加以解決，且為達目標願意主動承擔額外責任。
- A02自我管理：設立定義明確且實際可行的個人目標；對於及時完成任務展現高度進取、努力、承諾及負責任的行為。
- A03謹慎細心：對於任務的執行過程，能謹慎考量及處理所有細節，精確地檢視每個程序，並持續對其保持高度關注。
- A04持續學習：能夠展現自我提升的企圖心，利用且積極參與各種機會，學習任務所需的新知識與能力，並能有效應用在特定任務。
- A05好奇開放：容易受到複雜新穎的事物吸引，且易於接受新觀念的傾向。
- A06團隊意識：積極參與並支持團隊，能彼此鼓勵共同達成團隊目標。

說明與補充事項

● **建議擔任此職類/職業之學歷/經歷/或能力條件：**

- 大專以上資訊相關科系畢業，且具備2年以上資料處理相關工作經驗。

● **其他補充說明：**

- 資料庫技術：如 SQL、NoSQL 等。
- 程式語言：如 Java、C/C++、Python、.NET 等。
- 分散運算與儲存平台技術：如 Hadoop、Hive 等。
- 資料儲存分析平台：如 Hadoop、Splunk、Mongo DB、ELK 等。
- 分散運算與儲存技術：如 Hadoop、Hive 等。
- 資料處理、串流分析與技術應用：低結構化資料組織技術 (如 HTML、XML、JSON 等)、低結構化資料處理 (如文字、影音、影像與機器資料處理等)。
- 資料分析工具：如 SAS、SPSS、Splunk、R、Python 等。
- 資料處理語法：如 SQL、R、Splunk language 等。