

# 《資料處理》

一、為了有效衡量各部門的營運績效，單位主管或上層監督機關常訂出各單位部門的關鍵績效指標（KPI），訂出適當的KPI可以導引資源有效地運用。請以資訊部門為例，訂出適當的KPI。（20分）

<b>試題評析</b>	此題目屬於資訊管理應用題型。今年度考題，相對較不「技術」，以往常見的資料結構、資料庫、網路技術、程式語言等，幾乎一題都沒出現。今年度考題，對於資訊管理實務經驗較多的考生，相對較佔優勢。尤其前三題，實務應用的發揮，高過教科書上的標準答案。此題考的KPI，是介於資訊管理與人力資源績效管理的議題之間，這兩個科目都會教到。如果有企業實務工作經驗的考生，其實很容易在企業績效管理上遇到設定KPI這件事。因此，題目實際上並不難，但資料處理考這個，是有點文不對題。
-------------	--

**答：**

關鍵績效指標訂定的方式，一般分為由上而下（Top-down）和由下而上（Bottom-up）兩方面進行。前者指從企業總體策略出發，由上而下展開（cascading）各部門或員工的績效指標。後者則一般由各部門主管與部屬溝通，找出每位員工最合理的績效指標。

此外，訂定KPI時，須符合SMART原則：

- 1.明確的（Specific）：績效目標必須具體、清楚描述，不籠統。
- 2.可衡量的（Measurable）：可數量化，容易明確衡量是否達成或達成程度的。
- 3.可達成的（Achievable）：可實現的，不應訂立過高不可能達成、或過低無價值的目標。
- 4.相關的（Relevant）：與公司目標或工作明確關聯的。
- 5.有期限的（Time-bound）：有時間性的，明確定義目標應達成的期限。

以資訊部門為例，KPI例如：

- 1.於2016/12/31前，完成公司XX系統上線，提供至少OO名以上員工使用。
- 2.降低2016年資訊部門人力成本，目標為較前年減少20%。

二、政府機構將所擁有的大量資料公開已是大勢所趨，這些可被公開的資料稱為公開資料（open data）。請問公開資料具有什麼特性？這些公開資料可以產生什麼樣的價值？這產生出的價值可以怎麼衡量？（20分）

<b>試題評析</b>	開放資料（Open Data）並不是一個新名詞，現在網際網路相當普及，網路上可取得資料越來越多，除了企業風行的Big Data外，世界各國也由官方開始陸續推動Open data計劃。近幾年國內政府各單位也積極推動開放資料，雖然成效尚有相當成長空間，但開放資料儼然是電子化政府近年的新議題。但當然，開放資料由於是免費且容易取得的，維護資料的品質與個資法的把關也是在開放資料推行時責無旁貸的任務！
-------------	--

**答：**

（一）開放資料的主要特性為

- 1.開放（Open）：資料可免費自由被任何人使用。
- 2.利用與存取（Availability and Access）：資料容易使用、可重製、散佈。可被機器讀取（machine readable）的資料格式，且沒有使用與散佈的限制。
- 3.搜尋與發掘（Searchable and Discoverable）：資料容易搜尋取得。

（二）開放資料的價值：

- 1.促進透明化和民主控管（Transparency and Democratic control）
- 2.促進參與（Participation）
- 3.自我賦權（Self-Empowerment）
- 4.促成改善或有新的產品和服務（Improved or new private products and services）

- 5.促進創新 (Innovation)
- 6.提高政府服務效能 (Improved efficiency of government services)
- 7.改進政府服務效率 (Improved effectiveness of government services)
- 8.更有效的測量政府政策影響力 (Impact measurement of policies)
- 9.來自合併不同資訊來源之大量資料所產生的創新知識 (New knowledge from combined data sources and patterns in large data volumes)

(三)衡量開放資料的方式，可從四個構面來看：

- 1.背景與環境 (context/environment)：開放資料的提供背景。例如：國家級的政府開放資料，或是特定區塊 (例如：醫療健康、教育、交通等)。
- 2.資料 (data)：資料本身的特性與品質。例如：法律上、社會上的開放性、資料相關性、資料品質、資料類型等。
- 3.使用 (use)：這些開放資料如何被使用，以及可能的使用成果為何？例如：存取資料的方式、使用的目的、使用這些資料的活動等。即誰 (Who) 為何 (Why) 去使用什麼 (What)。
- 4.影響 (impact)：使用特定開放資料所得到的效益。效益可以依據社會性 (social)、環境 (environment)、政治/治理 (political/governance)、以及經濟/商業性 (economic/commercial) 幾個維度來看。

三、資訊系統開發前，需要了解與蒐集系統需求。若政府機關要開發或發包一套系統讓民眾使用，要如何儘可能地蒐集完整需求？(20分)

試題評析	需求蒐集是系統開發與分析前期的重要工作。可以從蒐集需求的前提、方法、工具幾個角度來回答。
考點命中	《高點·高上資料處理講義》第四回，唐箏編撰，頁57-58。

**答：**

需求蒐集是在系統開發的分析階段，首先必須先分析組織要利用資訊系統試著去解決的問題，包括定義問題、確定原因、確認解決方案，及配合系統解決方案所定義的資訊需求。通常是系統分析師角色的工作。尋找有該領域知識或系統經驗的系統分析師，對需求蒐集分析的正確、完整與成效，有著相當的影響。

一般需求蒐集方式：

- 1.需求擷取前，必須先了解系統潛在使用者以及可能之人機互動，以判斷需求蒐集的主要對象及角色。
- 2.接著，應蒐集欲電腦化之作業處理程序與其輸入資料內容、數量、格式、目標、規則與限制等。
- 3.常用的需求擷取方式有查閱文件、觀察、問卷、訪談、開會討論、與聯合開發 (Joint Application Development, JAD) 六種，這些方式可以單獨應用亦可相互搭配使用。

採用以上方式蒐集需求後，應將描述性需求轉換為系統性需求，透過需求塑模方式，例如：利用資料流程圖、UML等工具，將需求系統化詳述，並撰寫為需求規格書作為需求確認文件。

四、還不到幾年前，應用系統與資料檔案通常都儲存在個人的桌上電腦或筆電上 (本機端)，或所屬單位的伺服器上。現在資料與應用系統都漸漸地轉移到雲端，請分析比較資料與應用系統存在本機端與雲端的優劣。(20分)

試題評析	雲端應用的議題已經談了好幾年了。市場上的應用或服務也越來越多，考生對雲端應用有所了解，再系統化的分析優缺點即可。
考點命中	《高點·高上資料處理講義》第一回，唐箏編撰，頁80-84。

**答：**

		雲端	本機端
優點	成本	初期成本較低，例如：硬體伺服器、儲存成本。維護成本預算管理容易 (通常雲端服務將維護成本轉換為固	長期成本較低。以長期成本來說，較高的初期成本與後續維護費用，在相當之長期使用後，將很可能低

		雲端	本機端
		定，例如：每月支出）。	於雲端長期固定的租金支出。
	安全性	安全的雲端環境。許多雲端服務提供商都提供嚴謹符合安全法令規範的雲端環境，確保客戶資料不受威脅。	自行管理安全性。許多企業不希望資料暴露在網際網路之上，自建則可以容易限制系統或資料存取範圍。
	部署與擴展性	雲端服務通常依照需求部署所需資源、容易擴展，且需要較少的IT投入。	自行控制程度較高。較容易自行控制軟硬體部署時程，且企業對應用掌握度較高。
	使用者存取	行動工作者。容易遠端隨時隨地存取資料或服務。	網際網路連結性。不一定需要連結網際網路便可使用服務。
缺點	成本	長期成本高。雲端計費結構通常為租賃方式，長期租賃成本可能高於購買建置於本機端。且難以控制服務提供商的計費，如：價格調漲。	1.額外費用。例如：IT人力成本、軟硬體教育訓練費用、系統維護成本等。 2.硬體成本。本機需要自行購買建置伺服器、資料庫等。
	安全性	雲端資料與服務建置於服務提供商，服務提供商是否能確保資料安全性是關鍵。若非嚴格把關的提供商，則資料或應用可能有暴露於網際網路的風險。	需要專門規劃本機端應用或資料的安全性政策及方案，且必須自行考量系統的災害復原方案，例如：伺服器或資料庫損毀時的回復計劃。
	部署與擴展性	雲端服務通常較難達到企業客製化需求。需配合雲端服務商所提供的方案。	需要投入IT人力學習、部署、維護、與擴展系統。系統導入時程較長。
	使用者存取	雲端服務需要穩定且足夠的網路存取。如網路無法使用則雲端服務亦無法存取。	必須自行控制企業系統規格與頻寬，尤其是當同時使用人數眾多時。

五、一般資訊系統的功能與效能都是固定的，若要提升既有系統的功能，在軟硬體平台不變的前提下，都需要人為的介入去修改系統的原始碼，用更有效的演算法或資料結構以提升系統效能。但現在機器學習的方法可在不需要修改程式碼的情況下提升系統效能，請說明機器學習的原理並舉一例子解說。（20分）

試題評析	機器學習是偏計算機科學或資訊工程的領域，通常資料處理不會提及太多。但廣義的機器學習，其實包含了一些統計理論以及資料探勘，因此解答以資料探勘的角度來回答，考生們比較容易上手。
考點命中	《高點·高上資料處理講義》第三回，唐箏編撰，頁45-47。

**答：**

機器學習理論主要是設計和分析一些讓電腦可以自動「學習」的演算法。機器學習算法是一類從數據中自動分析獲得規律，並利用規律對未知數據進行預測的算法。機器學習常被歸納為人工智慧(AI)的科學，該領域的主要研究對象是人工智能，特別是如何在經驗學習中改善具體算法的性能。例如：類神經網路、決策樹、資料挖掘等，都是機器學習典型的一些技術。

舉例，透過Data Mining的Clustering技術，分析購物網站上相似類型的用戶會購買的商品。隨著資料量的增加，資料母體越大、以及得到的購買結果越多後，Clustering技術不需修改程式碼，但可做出更精確的資料叢集與行為分析。

人工智慧領域的機器學習，有更多更精細且複雜的機器學習理論，可在不需調整原始碼的情況下，餵入越大的資料量，做出越精確的決策，即「機器的學習」。

# 高點 · 高上

【版權所有，重製必究！】