

## 第四章 以知識為本之協同文件管理平台架構

本章主要採用框架演繹法，進行以電子檔案作業為研究主體的協同文件管理平台架構與流程分析。

### 4.1 以知識為本之電子檔案協同管理作業流程

前期研究經驗所述之結知識管理之文件作業流程（參見圖二-6），已具體表明電子檔案管理作業與其知識管理作為間之交互關係，本研究遽以此項研究成果為基礎，更進一步分析中央與各地區檔案保存機關間之「電子檔案管理作業流程」，如圖四-1所示。

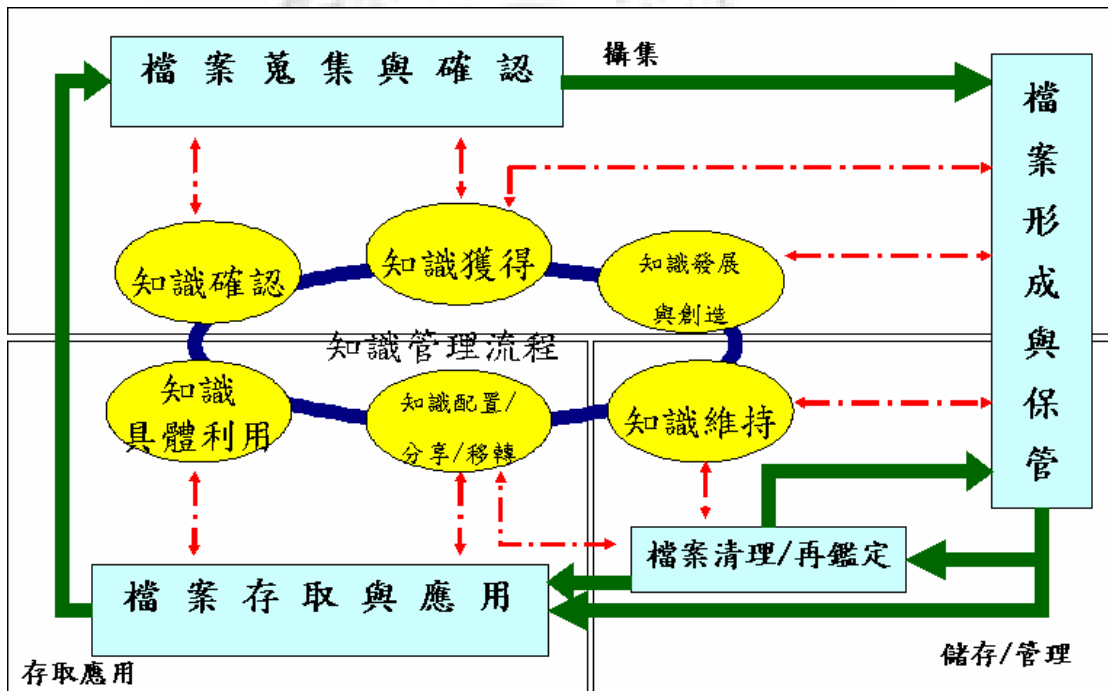


圖 四-1：以知識為本之電子檔案管理作業流程

資料來源：修改自(姜國輝、林誠謙等，民 93)

上述檔案管理作業流程與圖 二-6 的不同之處，在於上述作業流程對檔案清理作業進行更清楚的界定，檔案清理作業係為機關檔案作業與中央檔案作業最大的不同點，亦是唯一映射出檔案「永久保存」與「定期保存」相異特質之管理作業。本研究將檔案清理作業分為兩項子作業：（1）檔案清理/移轉子作業；（2）檔案再鑑定子作業，前項子作業係針對地方機關端之電子檔案，清除已屆保存年限，並無繼續保存價值之定期保存檔案，並對已達保存年限之永久保存性檔案進行移轉作業，進入中央檔案管理機關儲存；後項再鑑定子作業，則針對存於中央檔案主管機關之永久保存檔案，為使所保存檔案皆具有策略性意義，中央亦需定期為所謂「永久保存」性質之檔案，進行保存價值再鑑定工作，以因應當代政策與保存管理策略，並汰除已不符合現實所需之舊有永久保存檔案，持續對不斷擴充膨脹之國家檔案庫進行去蕪存菁之工作。

藉由此項管理作業流程，檔案中央主管機關與地區檔案管理機關便可以檔案知識管理流程為基礎，緊密結合彼此之文件與知識作業流程，構成電子檔案協同管理作業螺旋（如圖 四-2），以發揮電子檔案內涵知識之最大效益。

圖 四-2 即為電子檔案協同管理作業螺旋示意圖，居於作業螺旋底部的，為檔案知識管理作業流程，為組織所有文件管理工作共同的依歸，機關檔案管理作業即（內圈檔案管理作業）基於知識管理基礎為發展，經過蒐集、確認、形成階段，進入機關檔案保管系統典藏管理，對外發展適當的查詢介面以供使用者存取應用，對內亦定期進行檔案清理與移轉作業，以維持所保存檔案之價值與時效性；而永久保

存檔案經由移轉後進入國家檔案保存作業流程（外圍檔案管理作業），於徵集驗證後除進行長期典藏與應用之外，亦需如機關管理作業一般，定期進行檔案價值再鑑定工作。

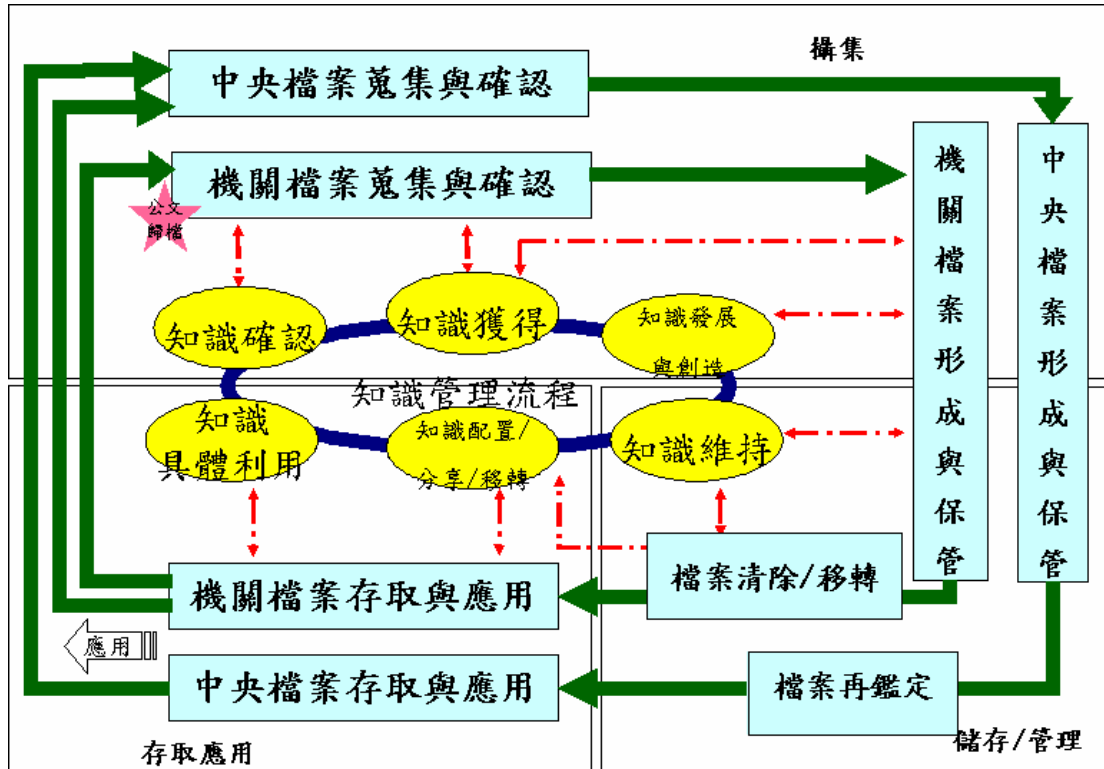


圖 四-2：電子檔案協同管理作業螺旋

## 4.2 文件資料管理之互通性框架

協同文件管理之目的即為解決大量知識文件之共享與儲存資源管理問題，而為使其於轉瞬萬變的科技環境中，仍能長期而穩定的運作，資料儲存與處理技術之變更管理相當重要。因此，知識文件保存所涵蓋之層次，不單僅是資料保存機制之實現而已，如何將資訊與知識及處理邏輯隨著文件一併留存，牽涉資訊與知識層次之相關邏輯甚

廣。本研究採用前述美國 NARA 暨 SDSC 提出之知識網格參考框架為基架，並深入展開如下，以竟知識文件分散儲存、協同管理之目的。

知識網格係以矩陣式組成架構，來設計一致性的管理機制，以協調儲存資源間的運作，俾使知識文件之知識、資訊與資料得以無礙地於各資源點間溝通傳遞，並正確地存取解譯。本節逕由知識網格為分析框架，更進一步分析以知識為基礎之文件長期保存系統應具之技術架構。

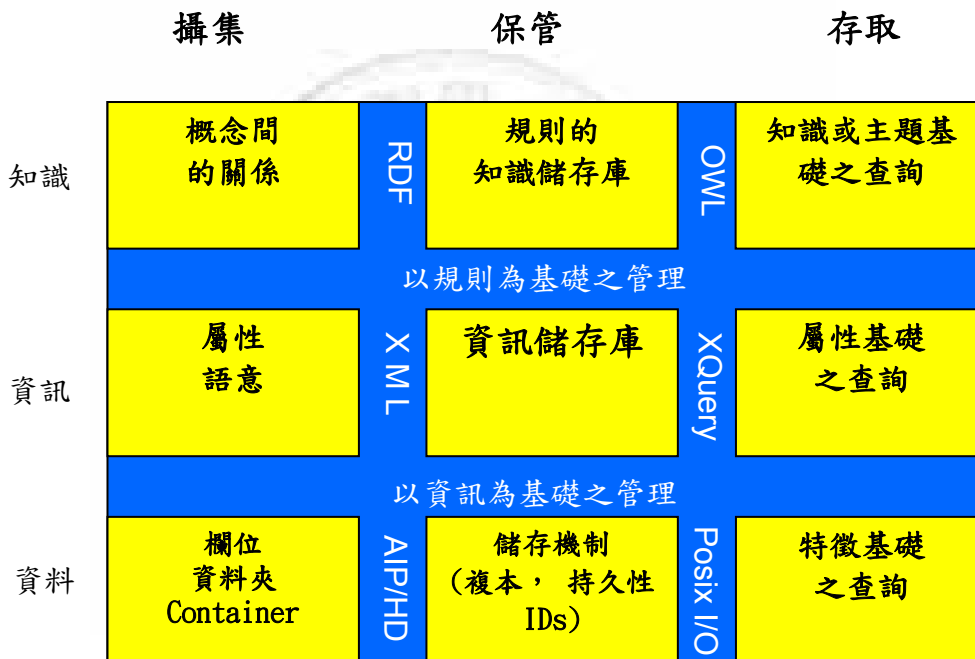


圖 四-3：知識網格

資料來源：Moore，2003c

由上圖之垂直軸表示此框架各層次之儲存元素：資料、資訊、知識。

- 資料(Data)是指檔案，亦即數位物件，如文件、影像檔等。一般由檔案或典藏系統所管理，並儲存於儲存庫 (storage

repository) 中。此間典藏的特點即是妥適留存其字元碼 (bits)。

- 資訊(Information)是指數位物件之屬性，可表現其文脈 (context)、來源 (Provenance) 或其他語意特徵，通常以後設資料的形式儲存於資料庫中；
- 知識(Knowledge)是指機關業務領域知識，以資訊屬性間之語意/邏輯關係、時間/程序關係、空間/地點關係或功能依存關係的形式表示，儲存於知識儲存庫中，由推理引擎提取使用之。

各層次間需注意資料、資訊與知識間之對應關係維護，也就是說，在知識和資訊間要訂定出從知識概念至屬性的對照表；在資訊和資料之間，能藉由資料控制系統，找出屬性對應之文件區塊儲存位置。

由檔案管理生命週期角度來看，整體以知識為基礎的長期保存檔案架構，需能建立一個貫穿攝集、儲存管理及使用三部分之間的軟體架構，相關之設計概念，簡述如下：

#### 4.2.1 攝集階段(Ingest)

此一階段，應著重的是組織數位物件之能力。

##### 資料層

資料的儲存與使用，重點在於異質儲存系統間之整合。於資料層次該注意的，是如何配置檔案與規劃其實體結構，以達成後續資料處理的效率。文件紀錄必須有唯一性識別碼，才能保持文件之一致性及

與後設資料進行對應。檔案儲存需將實際儲存管理與邏輯配置等作業獨立開來，以達成基礎技術獨立與位址獨立性。檔案之間的結構關係將透過後設資料屬性值表示，儲於資訊儲存庫（實則為機關內部之私用註冊機制）。

### 資訊層

資訊層次對應的，是建立與標示重要屬性與提取語意。於作法上即是 Tagging 的步驟。換言之，就是登錄長期保存所需之相關資訊：描述資訊、保存資訊、指引資訊、操作資訊等，並以後設資料的形式儲存資訊儲存庫中。

### 知識層

而知識攝集，係經由一系列 Mining 的內涵步驟對文件內容進行分析，提取文件知識，並建立跨概念間的對應。檔案的知識內容將以專門描述知識之後設資料（如 Topic Maps），著錄於知識儲存庫中。

跨越三個層次的，是一致性對應機制：如何將與某一文件相關的所有典藏性資訊與知識標籤，與其唯一識別碼對應，致使後續管理存取能產生一致性，此處由資料網格發展出之邏輯命名空間（Logic Name Space）機制即支援此一致性對應機制(Moore and Merzky, 2003)。

## 4.2.2 管理階段(Management)

此階段之任務為產生保存文件型式 (Archival Form)，並管理典藏文件及典藏性資訊之儲存狀況。面臨的挑戰是典藏紀錄與儲存硬體技術之間的獨立性，以及因應災難而生的損失問題。

故儲存管理之重點是實現平台獨立性，建立一個各層次間皆能獨立運作的儲存管理架構。換言之，即是可自我維持一致性的儲存管理機制。一致性的限制可能是時間上的，也可能是邏輯上的或結構上的。

各層次儲存庫間需定義一致的綱要，以利元素提取與交換，如：知識層的 XTM DTD，資訊層的 XML DTD，資料層的邏輯資料夾 (Logical Folders)、實體聚合(Physical Containers)和後設資料目錄 (Metadata CATalog，MCAT)。

### 資料層

複本(replicas)是提昇分散系統中檔案處理效率的一項作法，各地複本應永遠等價 (equivalence)，因此「複本間一致性的維持」與「Concurrency Control」是另一管理要點。

### 資訊層

於資訊層次對應的，是資訊元件的儲存管理，資訊儲存庫 (Information Repository)的設計、管理為此區塊重點。資訊儲存庫的目的是要建立一個基礎建設 (Infrastructure)，使得資料在後續存取可以達到通透性，也就是毋須知道物件名稱，即可查得所需資料。

參考數位典藏研究領域之領導組織 RLG 與美國長期典藏技術研發團隊 SDSC 之相關研究經驗 (RLG, 2002; Moore, 2004a), 本研究認為, 被典藏之元件資訊, 應包含以下幾項成分:

- (1). 內容物件: 所欲保存之標的物, 不僅為數位物件, 亦可為實體物件, 如檔案紙本。
- (2). 描述資訊: 提供可讓內容物件被了解之附加含義描述。如數位物件之資料型態、編碼格式、產生者、及產生時間等。於實體物件, 可存一系列分析之結果, 讓應用者更加了解其成分與結構。
- (3). 保存資訊: 以保存內容資訊之完整性。RLG 建議包含以下四類:
  - 參考(Reference): 提供使內容資訊可唯一鑑別之屬性值;
  - 來源(Provenance): 記載內容資訊之來源, 及操作歷史描述, 用以保證文件之信度;
  - 脈絡(Context): 描述內容資訊與外界環境之關係類型與關聯程度;
  - 不變性(Fixity): 保護內容資訊免於遭受未授權之修改。
- (4). 指引資訊: 輔助使用者搜尋所需資料, 應包含搜尋指引 (Finding Aids), 元件內容屬性及元件地點等之相關資訊。
- (5). 操作資訊: 隨著資訊技術的改變, 軟體解譯文件編碼之機制亦會改變, 因此, 將現有技術處理文件的方式留存下來亦是確保檔案長期解譯性之重要作為。此類資訊, Moore 將之稱為作用性(Behavioral)典藏資訊, 如檢視機制, 操作機制 (Moore, 2004)。



## 知識層

知識儲存庫(Knowledge Repository)是保存各邏輯系統 (Logic Systems) 間的關聯，使後續存取使用時具獨立性，毋須透過後設資料屬性名稱來組成資料庫。知識儲存之一致性限制，係指邏輯命名空間與分散資訊狀態的對應需一致。

為保持檔案完整性，不因災難及惡意存取而致檔案的損失，驗證授權等安全性機制與災難備原機制，亦為此階段之規劃重點。

### 4.2.3 使用階段(Access)

此階段的目標在於提供使用者與典藏系統之溝通介面，令使用者可以妥善應用蘊含知識，發揮典藏物價值。

在可預期的未來，電子檔案的數量會呈倍數成長。所以，檔案存取應用所面臨挑戰為「透通性」，讓使用者能夠從眾多的資料中方便簡易且快速地找到需要的文件。在透通性的部分，多方學者有相關文獻探討，由於 Moore 所考量的範圍最廣，本研究以 Moore 的理論為基礎，歸納重要的透通性，以利後續技術考量(Moore, 2003a)：

- 命名通透性：使用相關屬性值即能查詢需要的文件資訊，毋需預先知道其名稱。
- 位址通透性：毋需知道其實體路徑，即可存取所需物件。
- 典藏通透性：只要描述概念即可查詢使用，無須知道屬性名稱。
- 平台通透性：在不同儲存系統存取檔案，皆能順暢無礙。

- 編碼標準通透性：在不同平台上，正確展現數位物件之能力。
- 認證通透性：建立單點登入之環境，令後端於轉換儲存系統或是管理區塊時，使用者毋需再行登錄。

上述通透性，可經由標準的查詢機制達成。此套查詢機制，需於資料層次實現特徵為屬的資料搜尋；在資訊層，則可經由屬性值找出所需之檔案；於知識層，可以透過建構機關業務領域知識本體，完成以主題為主的跨領域查詢作業。

#### 4.2.4 區塊間介面設計概念

各區塊介面之技術選擇與開發，即是協助使用者存取通透性之實現。介面間共同需具備機制的為標準存取功能、標準傳輸協定、支援大量存取的匯出/入功能，以及邏輯命名空間之存取控制。

在攝集與儲存管理間，標準的標識（tag）機制（如 XML、RDF）的支援，可協助各層次元素之定義與傳輸。

在儲存管理與使用存取平台間，要支援以知識為基礎的查詢語言，如：知識層的智慧型查詢語言，資訊層的 EMCAT(Extensible Metadata Catalog)以及 SGL 產生器，資料層的運算網格等。此外，亦要能支援使用者權限資訊的互通，以實現單點登入服務功能。上述標準化操作集的形成，可以解決系統因技術變遷所導致之存取功能失效問題。

### 4.3 已知識為本的協同文件管理之互通性框架

採用知識網格，可達成儲存資源間的協調，基於此一致性的儲存管理技術架構之上，跨越組織與儲存資源的協同應用程式得以順利部署。而為協調組織間應用層面之業務流程與管理政策，便需參考應用 ebXML 等針對電子化協同作業所研擬之參考框架。而「知識網格」所構築之架構，恰可提供文件知識、資訊與資料間溝通之互通性參考。因此，本研究結合知識網格參考架構於 ebXML 開發框架中，提出「檔案知識保存與協同運作框架（Knowledge-based preservation and collaboration framework），簡稱為 KPC 框架」，如下圖所示。

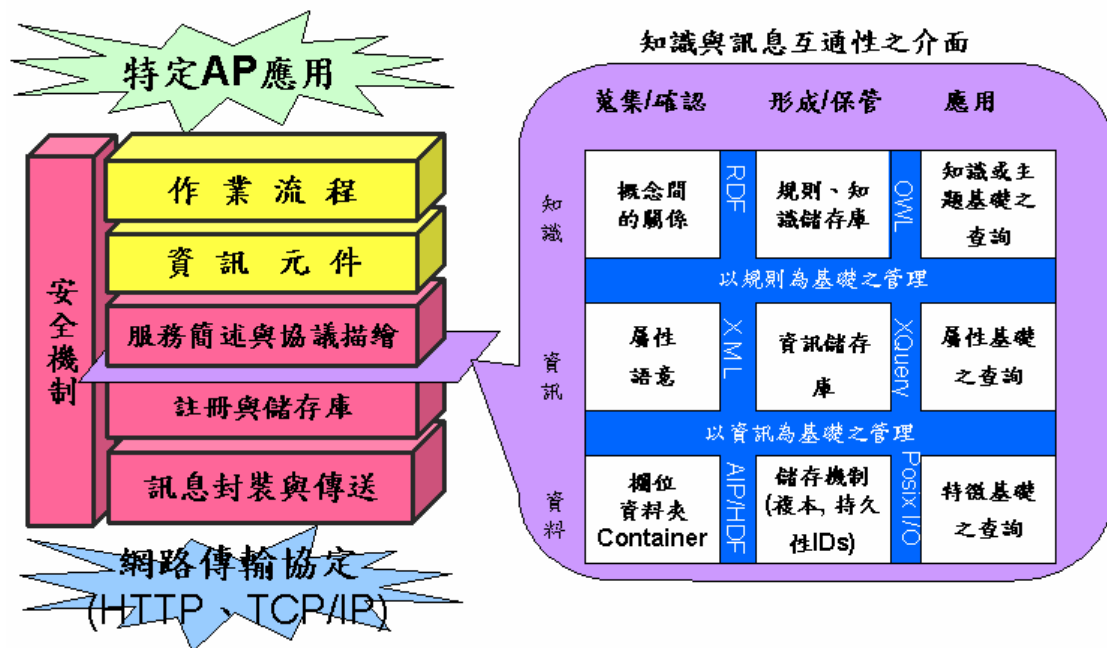


圖 四-4：檔案知識保存與協同運作框架：KPC Framework

註冊儲存庫之設計，可遵循知識網格所提供之建議。作業流程模型與資訊元件內所蘊含的知識與資訊，則對應於網格之各方格，進行蒐集與確認，並進而典藏於註冊儲存庫中；而網格架構中各方格間交錯銜接之介面，則與下層的訊息溝通機制相對應，已完成知識文件之分享交換。因此，兩個採用此協同運作框架所發展的組織，不僅能分別完成作業領域之知識描述與保留，以及下層訊息交換領域的溝通機

制，亦能將組織知識妥善地對應於儲存庫中，並以適當的方式分享於對方，於訊息互通性的基礎上實現組織間的知識互通性，如下圖所示。

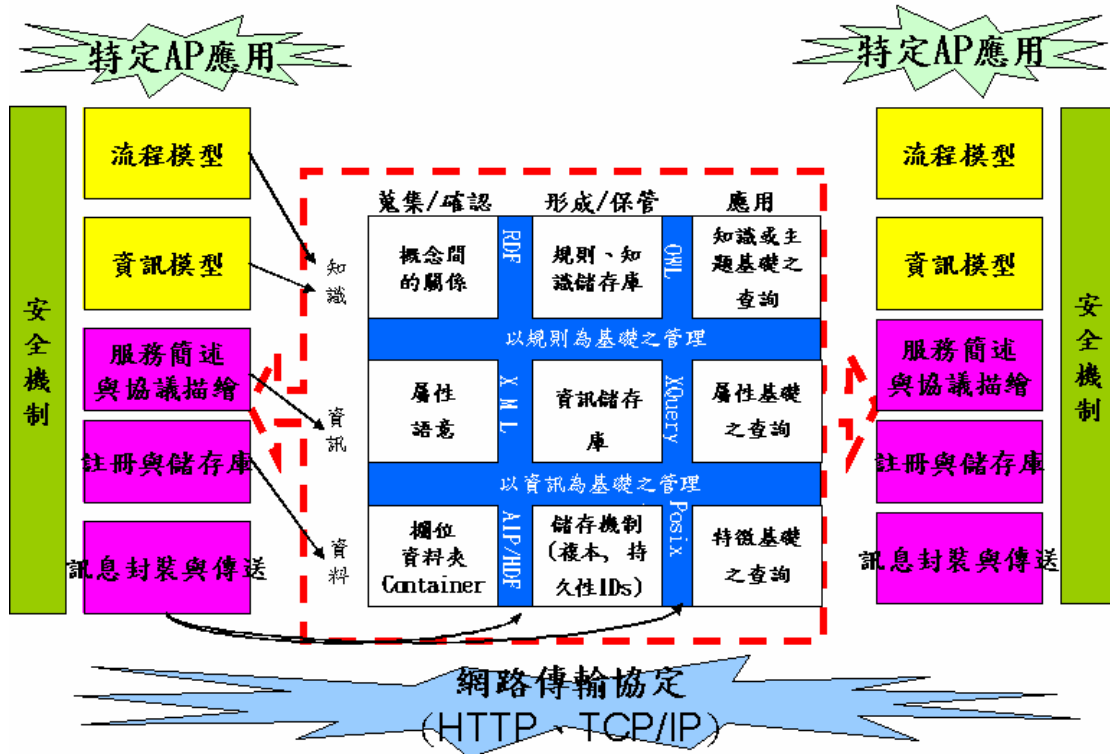


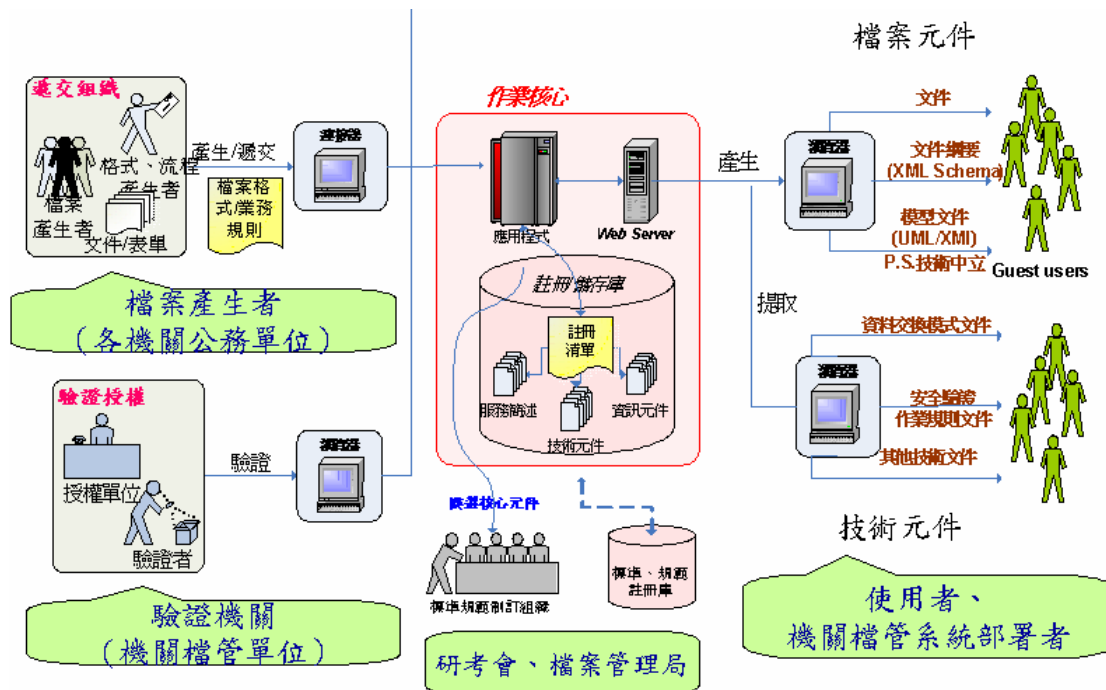
圖 四-5：具知識、資訊、資料互通性之架構

資料來源：本研究

KPC 框架結合電子商務應用服務框架與網格技術架構，使組織間進行協同作業應用時，其下層之溝通儲存機制亦能有其一致性架構。

#### 4.4 以知識為本之協同文件管理作業模型

本研究以上述 KPC 框架為發展核心，提出一以知識為基礎之協同文件管理作業模型如下，並於後續小節中說明其運作流程。



攝集

儲存管理

存取應用

圖 四-6：以知識為基礎之協同文件管理作業模型

資料來源：修改自 Najmi, Berthomieu, and Bourge, 2004

以下將對本模型之運作模式進行簡述：

### (1). 核心系統

平台架構之核心即為圖中央之註冊/儲存庫，茲用以註冊並儲存管理各成員之技術元件清單、核心資訊元件清單以及服務簡述檔，這些元件之產生皆遵循作業流程規範制訂組織（如 UN/CEFACT）所規定之描述規範，並於收錄儲存前自動進行完整性驗證。註冊機制與內部架構，皆遵循 UN/CEFACT 等方法學制訂組織所研擬之標準建立。

### (2). 文件格式/業務規則之驗證

此架構的特點即是：文件格式與業務規則的描述係開放性的。因此，當組織因變革而產生流程改善，或因資訊技術改變而產生資訊處

理程序改變時，皆可依規範製作流程或文件綱目，提交於核心系統等候驗證，如圖左上之業務遞交動作；系統接獲遞交項目後，將通知驗證單位（圖左下）進行驗證，驗證者根據自身之專業與技術環境，評估所提交之綱目，然後提出接受或駁回之裁示。

其後，驗證者可視新綱目之通用性，提交業務流程方法學組織進行驗證，等候成為通用標準。

### **(3). 應用程式部署**

當某機關決定建置並部署屬於自己且符合規格的應用程式，可透過核心系統應用程式，連上核心註冊儲存庫，下載相關業務及技術規範（如：相容之文件格式、處理程序模型、資料交換規則文件等）。取得規範後，即可建置自己系統的應用程式。

### **(4). 知識文件管理與服務能力描述**

各機關建置符合規範的保存系統之後，便可將主管之文件載入儲存資源中，並擁有完整的管理權，包含內容驗證權。典藏文件之內容與格式需根據核心規範驗證通過。各機關並製作「服務能力描述」清單，內容為本身所提供的服務能力與主管的文件清單，註冊並上載於核心系統中，以供可能的合作伙伴查詢。

### **(5). 知識文件查詢與應用**

由於核心註冊儲存庫載有所有成員之服務能力清單與所管理文件之清單，任何有查詢權限之使用者，皆能經由網際網路連上此核心系統進行查詢，獲取擁有此類文件之機關服務簡述檔，後續與文件主管機關協商，進行知識文件之應用服務申請並執行。

若所欲存取之文件其時日已久，本地端無格式解譯工具，則可採用兩種作法：(i)逕行對核心系統發出服務能力查詢指令，取得可提供格式解譯服務的機關服務能力描述檔，發出服務與協同合作需求；(ii)查詢提取著錄此文件格式之綱目資訊元件，再進行解譯服務。

#### 4.5 協同合作模型之採用建議

資料與知識密集(Data/Knowledge intensive)的應用服務牽涉大量的運算資源與儲存資源，因此，本研究將 Federation 合作模式採用於跨組織的資源與技術能力協調，以實現兼顧共享與自治的資源協調需求。

由 Federation 文獻探討中可看出，組織間資訊交換的頻率及節點管理結構為 Federation 運作模式之重要影響因素。本研究嘗試更清楚地畫分 Federation 類型與運作模式，故由「節點管理結構」與「資料/資訊分享頻率」兩大維度，構成四種合作類型，並重新分派上述之十種運作模式於此四項協作類型中：

- 節點管理結構：對等/層級。

在對等結構下，各節點間的角色關係彼此平等，沒有主從之分，資料交換之工作由連結節點相互協調；而採層級結構之管理節點有主/從之分，由主節點統籌管理資料同步與交換工作與時間表。

- 資料與資訊分享頻率：同步/交換。

組織間的資料交換頻率視其合作目的與協調程度，而有不同的交換頻率，如資料密集的計算節點間的合作，需要緊密的工作結合與高頻率的資料交換，定期的資料同步化交換是相當必須的；至於交換需求不大之兩節點間，則可於有交換需要時再進行資料交換即可。

本研究將「同步類型」定義為為節點主動或定期將分享資料/資訊同步化；而將「交換類型」定義為視應用所需分享資料/資訊。

表 四-1：Federation 合作類型-以節點關係結構/分享頻率為分類維度

關係結構 \ 頻率	分享	高（同步）	低（交換）
對等		3. 資料複本 4. 使用者與資料複本 5. 目錄複本	1. 偶而交換 2. 資源互動 10. 自由浮動型
階層		6. 主從型 7. 雪片型 8. 典藏型	9. 游牧型

基於上述互通性架構為基礎，組織協作模型之採用相當有彈性，在一組 Federation 中，各組織可以依內部需求與管理結構，決定其資源間協同模式；而組織間之協同，則以聯盟成員的共識來擬定其組織間 Federation 模式。因此，在分析的時候，必須釐清 Federation 層次，才得以洞悉需求，建構出最適用之運作模型。

由組織間合作的角度來分析，具有知識文件長期保存需求的組織，多半為需長期永續經營的非營利組織、研究組織或政府機關與國際組織等，如：國家型博物館、圖書館、檔案館、國際遺產保護組織、科學研究中心及數位典藏機構等。然而聯盟中各單位之合作緊密度並不一致，拿國家檔案館、某地方檔案管理機關與國史館而言，前兩者彼此為上下屬關係，且業務流程互相銜接，其合作緊密度便遠大於後兩者間之合作關係，因此各節點間應視彼此合作之緊密度與頻率不同，而採用不同的 Federation 模式，故本研究建議如下：

當成員彼此作業協調需相當緊密，且資料分享安全性無虞，則建



議採用對等同步的合作模式。如科學性研究組織之合作研究計畫等。

當成員彼此階層合作關係不明顯，且彼此不牽涉營利目的，但許多文件具有相當高的國家/組織機密性，對等且交換的 Federation 模式較適用上述合作情形。跨國際的人文社會性研究合作計畫多屬此類。

隸屬於同一管理政策下，文件交換之機密需求非其重點，且組織間多半有從屬關係者，採階層同步的合作模式為宜。如國家內各機關的協同合作。

階層交換型合作多半用於一般使用者資源與主管機關之協調，由於使用者為動態的連結所需的組織，合作時間短，所需的服務變化性大而非例行性事務，故採用階層交換型合作為宜。

