

# GKS - 群組文件與知識典藏系統

鄭心怡

林宜華

何建明

中央研究院資訊科學研究所

臺北市南港區研究院路 2 段 128 號

E-Mail: {hycheng, shlin, hoho@iis.sinica.edu.tw}

## Abstract

隨著網際網路的普及，網路上無數的使用者與文件資訊透過 Web 網路連結成一個世界上最大的知識網路，面對這個無遠弗屆、不斷成長的知識網路，已經有 Google 等搜尋引擎，協助使用者拓展未知的知識領域。但是，如何從這些資源中擷取出有用的資訊，並整合群組成員共同生產的知識，進而加以有效管理、應用，對於知識工作者而言，卻是一個新的挑戰。其中問題包括：而來自網路上大量不同來源與格式的文件，造成使用者典藏與管理上的不便；而組織內的各成員更有不同的作業環境，使得成員間的知識分享也有一定的困難。為因應這些問題，我們建立了“群組知識共享、搜尋與管理系統”(Group Knowledge System, 簡稱 GKS)，為合作群組的成員提供知識擷取、整合、典藏、檢索、分享的工作環境。成員可透過瀏覽器來使用系統所提供的服務。目錄管理系統提供目錄編輯，建構各成員本身或組織的知識架構；利用群組管理，可依照組織內成員各自的責任與專業，設定權限來共同編輯、彙整組織的整體知識架構 (Global Knowledge Ontology)。本系統能索引與檢索目前網路上最常使用的文件格式 (HTML、TXT、MS Office、EML、PDF 等)，並支援各種資訊來源 (網站、資料庫、內部檔案等)，讓不同來源的不同資訊能在同一目錄知識架構下管理。利用自動分類學習的技術，系統能從目錄知識架構所蒐集的文件資訊學習分類知識。以學習到的分類知識為基礎，系統可到搜尋引擎 (例如：Google) 檢索分類相關資訊；或經由文件內的連結文字和分類知識的相似度，自動蒐集網路上的相關的資訊。利

用這些方式，系統可自動擴充知識版圖。所以，本系統結合網路上的使用者去蒐集網路上無限的資訊，加以整合。同時，本系統也提供網路儲存服務 (Web storage service)，包括目錄和使用 PC 檔案系統同步的機制，以方便個人或企業內部的文件資訊也能整合於本系統，建構組織內完整的文件和知識典藏系統。上述這些功能，再加上系統所提供的檢索系統 (可檢索各式文件、資料庫、網站、metadata 等)，配合群組管理對成員權限之設定，組織可快速分享知識給適合的成員。因此，本系統所提供的平台，方便成員共同建立出組織的知識典藏體系，並與組織成員分享網路的知識，以滿足組織和成員在知識經濟時代對資訊和知識的需求。系統開發人員也可以在這套系統之上，進一步開發各種工作流程，以進一步簡化文件的管理功能。這套系統目前網址在 <http://gks.iis.sinica.edu.tw/>。

## 1. 簡介

隨著網際網路的普及，網路上無數的使用者與文件資料透過 Web 網路連結成一個世界上最大的知識網路。面對這個無遠弗屆、不斷成長的知識網路，使用者可以利用搜尋引擎 (例如：Google [1]) 獲得所需資料，而如何從這些龐雜的資料中轉換、整理出真正有用的資訊，並經過分類歸納，獲取規則進而建立知識，已成為每個知識工作者的重要課題。而在組織中，各成員的知識要如何分享、傳遞，並整合群組成員共同生產的知識，進而加以有效管理和應用，更是一個新的挑戰。以九年一貫教學資源網站為例，為了將九年一貫國民教育課程資源彙整成一個完整的教

學資源中心，由各學習領域負責人共同決定資源中心的整體結構與各學習領域的分工方式，再由負責人與各學習領域課程小組成員建構領域知識，種子教師將資源網站的使用普及到全國各學校的教師和學生，甚至於一般大眾。所以，在系統平台的設計中，除了提供組織內合作群組的成員一個知識擷取、典藏、檢索的工作環境，如何讓資源在上述成員所組成的階層式虛擬社會的模型中達成知識的整合與分享更是相形重要的。

**Group Knowledge System (GKS)** 即是以上述目的為出發點所實作的系統。以目錄架構為基礎，來管理各項網路資源（來自文件、資料庫、網頁、電子郵件等）。同時，也提供階層式架構的群組和成員管理，讓網路上的使用者所形成的階層架構，共同管理網路資源。

在下面的論文中，我們先探討和 **GKS** 相關的研究，接下來在第三、四、五節剖析系統設計和實際的管理與應用。最後提出結論。

## 2. 相關研究

Thomas H. Davenport [5] 曾指出，典型的知識管理工作項目包含了：

- 知識擷取，即創造文件並把文件輸入電腦系統。
- 透過編輯、組合和整理為知識增加價值。
- 發展知識的分類方法。
- 散播知識之資訊科技基礎建設。
- 教導員工創新、分享與使用知識。

而這些工作往往需要投入大量的金錢與勞力。**GKS** 就是為了因應這些需求，提供組織內的成員一個知識典藏與分享的工作平台，將電腦、網路、和人這三種系統需緊密配合，以建立一套良好的知識管理環境，提供知識解答做為決策參考。

和 **GKS** 相似的系統則有 **Zope**。**Zope** [6] 是一個開放性原始碼應用程式伺服器 (open source application server)，可協助程式設計人員建置內容網站、入口網站、互動式多媒體

網站等等，提供團隊共同開發、管理動態 web-based 商務應用程式的環境。這個網站出版系統 (Web Publishing System) 包含了網頁伺服器、物件導向式資料庫、內容管理器 (content manager)、網站管理軟體 (site manager)，並提供安全機制。唯該系統在中文支援上仍待加強。

## 3. 系統設計

接下來我們將探討如何根據需求規劃 **GKS** 系統、資料庫綱要，及系統架構的設計。

### 3.1 系統規劃

為了符合前述需求，提供使用者能有效建立知識體系與進行知識分享的系統，我們做出以下的系統規劃。並且，考慮到使用者可能使用不同操作平台，**GKS** 提供 **Web-based services**，亦即使用者可透過瀏覽器來使用系統所提供的服務，如下列所示：

- **目錄和物件資源管理 (Directory and Resource Management)**：本系統將使用者的典藏資源分為目錄與物件兩大類，提供階層式目錄管理機制來管理使用者的典藏資料。目錄所表達的概念是大量資源的組織架構和分類，目錄和目錄間存在多對多的關係；而使用者的每一筆典藏資料都視為一個物件，目錄和物件間存在多對多的關係。使用者可以利用本系統提供目錄和物件資源管理工具，建構各成員本身或組織的知識架構，並以目錄作為分工的基本單位。
- **目錄搜尋與全文檢索 (Directory and Resource Full-text Search)**：使用者可透過目錄搜尋與全文檢索尋找所需的目錄與物件資源，使資源的取得更有效率。本系統能檢索目前網路上最常使用的文件格式，例如：**HTML**、**TXT**、**MS Office documents (DOC, PPT, XLS)**、**EML (E-Mail)**、**PDF** 和 **PS** 等，並支援各種資訊來源（網站、資料庫、內部檔案等），讓不同來源的不同資訊能在同一目錄知識架構下管理。

- 群組和成員管理 (Group and Member Management)**：在組織的知識分享過程中，知識要如何在組織中有效地傳遞，組織成員所組成的階層式虛擬社會的模型是非常重要的。利用群組、成員、角色所組成的階層式虛擬社會的機制，各組織部門間可共同建構、管理知識架構並進行溝通，以利長久經營。本系統可依照組織內成員各自的責任與專業，設定權限來共同編輯、彙整組織的整體知識架構 (global knowledge ontology)。設計如下：
- 群組**：群組包含多個成員，通常是為了某個功能或目的所設定，讓群組內的所有成員能共同享有權力，來完成此目的。群組間的成員允許重疊，亦即各成員的群組歸屬應以功能來區分；例如：系統管理群組的成員有權限管理系統；數學教材審查群組的成員可以審查此領域的教材；教材編輯群組的成員可以新增 (上傳)、修改、刪除、瀏覽和閱讀教材。系統並提供管理機制，使用者可以申請群組，系統管理員審核成功後，此使用者成為此群組之擁有者 (owner)，並有權力指派此群組之多個管理員 (manager)。一般使用者可申請成為此群組之會員 (member)，以共享此群組所管理的資源，群組管理者即可分工審查成員之申請。GKS 提供系統管理員可以管理所有群組品質，若群組的品質低落或行為逾越系統規範，系統管理員可以刪除該群組。
- 成員**：成員可參加各種群組並在各群組內擔任不同角色 (需通過群組管理者之審核)，以獲得不同的權力，做不同之資源貢獻和分享。
- 角色**：成員在群組中所扮演的角色，就是成員在該群組中所擁有的權限，例如：系統管理員、群組管理者、一般成員等等。在系統中，角色可依各群組的功能彈性和權限管理進行新增、修改或刪除。例如：在教材的審查流程中，新增審查員 (reviewer) 來審查教材，以避免和 managers 及 members 的角色混淆。
- 網路儲存服務 (Web storage service)**：同時，本系統也提供以 Web 為存取方式的儲存服務。包括目錄和使用者個人電腦檔案系統同步的機制，以方便個人或組織內部的文件資訊也能整合於本系統，建構組織內完整的文件和知識典藏系統；並且透過瀏覽器與系統的權限控管安全而便利地存取資訊。
- 目錄分類學習和自動分類 (Classification Learning and Automatic Classification)**：本系統並可利用自動分類學習的技術，從目錄知識架構所蒐集的文件資訊學習分類知識 [2][3][4]。以學習到的分類知識為基礎，系統可到搜尋引擎 (例如：Google) 檢索分類相關資訊；或經由文件內的連結文字和分類知識的相似度，自動蒐集網路上的相關的資訊。運用這些方式，系統可協助成員自動擴充知識版圖。所以，本系統結合網路上的使用者去蒐集網路上無限的資訊，加以整合。

### 3.2 資料庫綱要

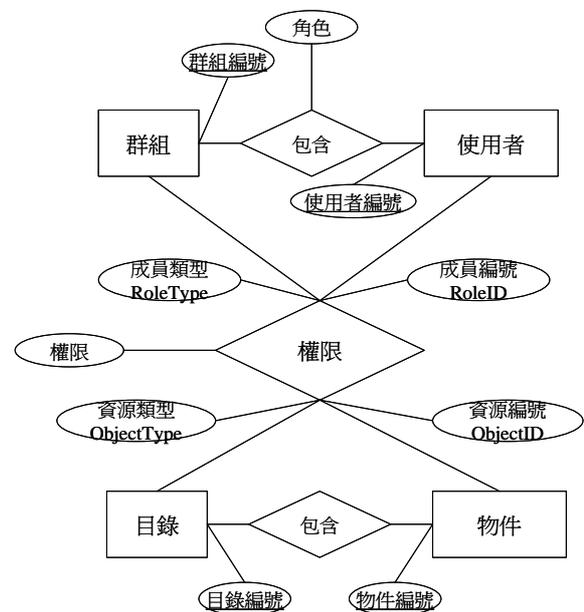


Figure 1. GKS: Entity-Relational Model

為了讓上述的階層式虛擬社會的模型能有效進行知識的整合與分享，使用者、群組與權限控管的機制應有更詳盡的規劃。由於一個

群組擁有多個使用者，每個使用者也可加入一個以上的群組，故群組與使用者間的關聯應為多對多，使用者在群組中擔任的角色應紀錄於該關聯中；而各群組與各使用者可對不同的目錄與物件資源擁有不同的權限，另建立一關聯資料表紀錄之。Figure 1 為相關的資料庫綱要之實體關係模組 (Entity-Relational Model)：

### 3.3 系統架構

GKS 架構在資料庫系統和檔案系統上，系統架構如 Figure 2 所示，從下層到上層描述如下：

- 權限管理：控制使用者、成員和管理者存取資源之權限。
- 群組工具：提供群組管理。
- 目錄和物件資源管理：提供一般化的目錄和資源管理。
- 目錄瀏覽搜尋和全文檢索：檢索系統資源 (目錄、文件、群組和成員)。
- 使用者工具：提供單一登錄和登入的窗口。
- 其他功能與工具：包括自動學習分類知識和自動分類系統、關鍵詞分析管理暨關聯詞和新詞學習系統等。

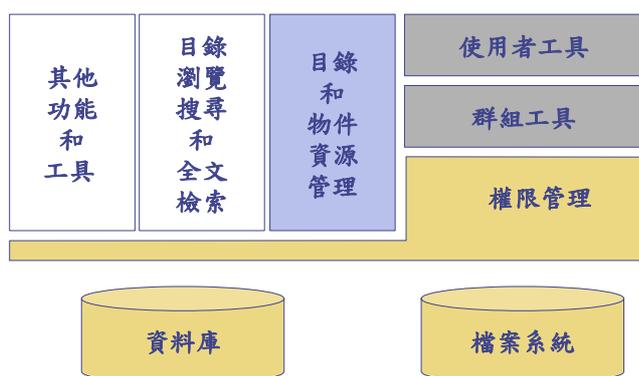


Figure 2. 系統架構

### 3.4 系統管理與應用

根據系統規劃和需求分析，我們定義出下列的系統物件和服務。

#### 3.4.1 群組與使用者、物件

- 使用者：一個使用者可加入多個群組，本系統成員均屬於一般使用者群組。系統管理者在評量使用者可以從以下幾點考量：  
使用者身分：使用者本身在許多群組中擔任管理員等重要角色，並且這些群組也有很好的品質。

使用者的知識架構：使用者本身擁有或管理的知識架構有很好的品質，像是目錄數目、物件數目、目錄與物件被瀏覽次數等等都可以作為參考的依據。

- 群組：分為系統群組與使用者自建群組，使用者自建群組由群組內的管理者管理，而群組由系統管理者負責管理與審查。由於只有群組管理者可檢視該群組成員。目前 GKS 系統定義的基本群組如下：

**Guests**：未登入系統的瀏覽者均給予此群組權限。

**Users**：GKS 系統成員均屬於此群組。

**Administrators**：系統管理者。

使用者自建群組：由使用者向系統管理者申請。系統管理者在評量群組的品質可以從以下幾點考量：

群組成員數量。

群組權限：群組在許多使用者的知識架構中都被賦予權限，亦即這個群組被系統中的多數成員所信任。

- 群組內的角色：一個群組可擁有多個使用者，其中只有一個群組擁有者，但可有數個管理者，其餘使用者在群組中也可擁有不同的角色。群組管理者可檢視群組成員內容與各成員權限；未加入群組的使用者或群組成員僅能檢視群組名稱。

#### 3.4.2 網路儲存服務 (Web storage service)

每個系統成員在在伺服器檔案系統中擁有一個檔案典藏的目錄，該目錄的路徑紀錄在資料庫中的使用者資料裡，使用者上傳的文件後，程式會讀去這個路徑，配合物件編號將

文件儲存於伺服器檔案系統中對應的目錄下。基於系統延展性(Scalability)考量，系統管理者可依系統擴充需求增加儲存系統(Storage Server)，規劃好需要變更檔案典藏目錄的成員，並修改資料庫中系統成員的檔案典藏目錄路徑與移動該成員的檔案典藏的目錄，即可完成儲存系統的擴充。

## 4. 系統實現

以下就各系統模組進行介紹：

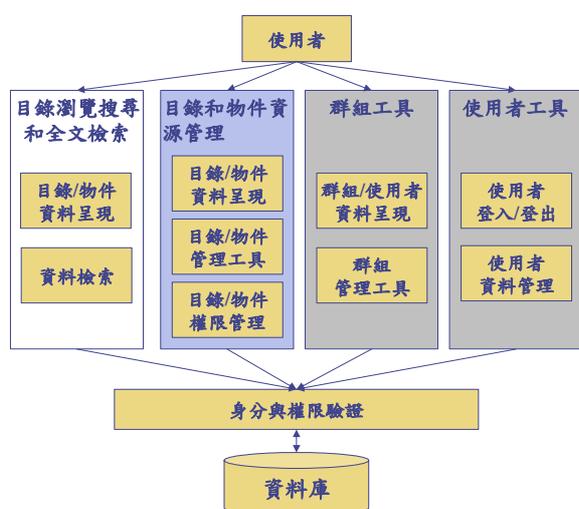


Figure 3. 系統模組

### 4.1 使用者工具

#### 4.1.1 使用者登入/登出

在使用者將登入資料從使用者端瀏覽器傳回後，檢查帳號密碼的正確性，確認無誤後將登入資訊寫入使用者端的 cookie，並將使用者多項資訊進行編碼存入資料庫，接下來的身分確認都以 cookie 中的這個編碼來進行確認。系統允許在同一時間內一個 IP 同時登入兩個以上的不同帳號，或者一個帳號可同時提供多人在不同 IP 使用。為了提供單一登入的機制，系統在寫入使用者端的 cookie 時，會設定 cookie 的網域名稱 (domain)，讓使用者在登入一次後便可以登入使用所有合作網站或服務。

### 4.1.2 使用者資料管理

GKS 提供使用者以下工具：

- 個人資料編輯管理
- 刪除帳號
- 密碼修改
- 所屬群組編輯

## 4.2 群組工具

### 4.2.1 群組/使用者資料呈現

呈現群組名稱、描述與屬於本群組的使用者帳號、角色等資訊。由群組資料呈現、群組資料管理等程式共用。依使用者不同的身份與權限與程式參數，呈現不同的編輯介面。例如，在群組資料管理中，若使用者為群組管理者，則 GKS 會另外呈現群組管理工具。

### 4.2.2 群組管理工具

使用者可申請群組以做更有彈性的權限管理。經系統管理者審核後，此使用者就成為群組擁有者，可利用群組管理，審核一般使用者申請成為群組成員、設定群組內使用者角色、新增使用者。

## 4.3 目錄和物件資源管理

### 4.3.1 目錄/物件資料呈現

呈現目錄名稱、描述與該目錄中的目錄、物件等資訊。本模組與目錄瀏覽搜尋和全文檢索共用。版面呈現資訊可依使用者設定而有不同的呈現。

### 4.3.2 目錄/物件管理工具

提供目錄與物件資料編輯工具，包括目錄與物件資料的新增、修改、刪除、移動，並提供檔案上傳、檔案下載、與瀏覽器中的書籤同步等協助系統目錄和使用者個人電腦檔案系統同步的功能。

### 4.3.3 目錄/物件權限管理

管理者可利用目錄與物件權限管理工具編輯所有使用者與群組對該目錄或物件的存取權限，並可設定子目錄及物件的權限繼承父目錄的權限。為了方便權限管理，GKS 也提供 ACL (Access Control List) 的設定，讓使用者定義一些 access controls，作為不同權限的管理。例如，可在 AC\_Reader 定義一組成員和群組，這些人只有 Read 的權限；同樣，AC\_FullControl 可定義出一組完全控制的權限。這些 access controls 可以在任何權限管理使用時共用，例如我們正在發展中的 workflow system。

#### 4.4 目錄瀏覽搜尋和全文檢索

目錄瀏覽共用上述的目錄/物件資料呈現模組。全文檢索提供使用者輸入的搜尋條件來檢索符合的目錄與物件資源，系統提供欄位檢索和 AND/OR/NOT 等邏輯運算。

### 5. 系統應用

以九年一貫網站建置為例，在教育部推動「資訊教育總藍圖」案中，預定建構生命教育學習網、健康醫學學習網、自然生態學習網、歷史文化學習網、人文藝術學習網、科技教育學習網等六大學習網，同時為協助中小學教師透過這些網站進修或擷取教學資源。因此，整個計畫的執行需透過階層式的傳播，經由種子教師將六大學習網的使用普及到全國各學校的教師和學生，甚至於一般大眾。若以本系統為基礎平台，應先由六大學習網負責人協調討論整體的知識架構與各學習領域的分工方式，將整體知識架構下的各學習領域交由各學習網負責人與工作小組負責建置，種子教師則是建置初期的網站使用者。

在這樣的成員架構下，在系統建置初期應建立六個群組，以各學習網負責人為群組擁有者，各工作小組為群組成員，並可從中挑選適當成員為群組管理員，種子教師則為系統一般使用者（系統群組）。各群組擁有各學習領域的編輯權限，種子教師則視其分工在不

同的知識架構中給予檢視或編輯的權限，或是系統一般使用者均給予檢視的權限。

### 6. 結論

這篇論文針對組織面對網際網路中大量不同來源文件資源、組織內各成員不同的作業環境而發生的知識分享問題，發展了「群組知識共享、搜尋與管理系統」(Group Knowledge System)。

本系統提供 Web-based services，亦即使用者可透過瀏覽器來使用系統所提供的服務，以此解決組織成員可能使用不同操作平台的問題；以階層式目錄架構為基礎，來管理各項網路資源(文件、資料庫、網頁、電子郵件等)；同時提供階層式架構的群組和成員管理，讓網路上的使用者所形成的階層架構，共同管理網路資源，以達知識分享之目的。組織可以以本系統為合作群組的成員提供知識擷取、整合、典藏、檢索、分享的工作環境，建構組織成員與組織整體的知識架構。

### 7. 參考文獻

- [1] Google search engine, <http://www.google.com/>.
- [2] Lin, S. H., Chen, M. C., Ho, J. M., and Huang, Y. M., "ACIRD: Intelligent Internet Documents Organization and Retrieval," IEEE Transactions on Knowledge and Data Engineering, Vol. 14, NO. 3, May/Jun, 2002.
- [3] Lin, S. H., Shih, C. S., Chen, M. C., Ho, J. M., Ko, M. T., and Huang, Y. M., "Extracting Classification Knowledge of Internet Documents: A Semantics Approach," Proceedings of the 21st International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'98), Melbourne, Australia, August 24 - 28, 1998.
- [4] Lin, S. H. and Huang, Y. M., "An Efficient Inductive Learning Method for OODB Using Entropy, IEEE Transactions on

Knowledge and Data Engineering, Vol. 8,  
NO. 6, December 1996.

<http://www.bus.utexas.edu/kman/kmprin.htm>.

[5] Thomas H. Davenport, "Some Principles of  
Knowledge Management",

[6] Zope, <http://www.zope.org/>.