

行政院國家科學委員會專題研究計畫結案報告

數位博物館計畫：玄奘西域行—子計劃三

數位博物館資訊技術之研發

Information Technologies of Digital Museum

計畫編號：NSC89-2750-P-002-004

執行期限：89年1月1日至89年12月31日

主持人：許清琦¹

共同主持人：歐陽彥正

參與人員：蕭文宏、巫有欽、章秉純

執行單位：國立台灣大學資訊工程學研究所

E-mail¹: cchsu@csie.ntu.edu.tw

一、子計畫三「數位博物館資訊技術之研發」成果說明

子計畫三「數位博物館資訊技術之研發」在各組的配合下，完成相關系統程式與應用程式設計。本子計畫旨在發揮資訊技術所帶來的進步與效率，建構一個研究或學習佛學知識的數位化系統環境，提供給一般民眾輕鬆且友善的佛學研究資訊系統。第一年度子計畫三主要有兩項任務，第一是配合子計畫二之需求，研發使用環境及資訊呈現軟體。第二是配合子計畫一之需求，研發高效率的資訊檢索軟體及內容管理機制。根據子計畫一所收集、組織、整理的內容，如佛學書目、佛教經文中的《大正新脩大藏經》、佛教辭典中的丁福保《佛學大辭典》、佛教典故中的《玄奘大師的故事》、《大唐大慈恩寺三藏法師傳》及《大唐西域記》等相關資料，研發設計出『佛學書目檢索系統』、『大正藏多版本經文檢索系統』、『佛學辭典檢索系統』及『玄奘傳·西域行之多文獻比對與展示介面』。

(一)、完成之工作目標

一個數位博物館設立的成功與否以及是否受歡迎，除了其內容是否深具典藏與傳承的意義、以及網頁設計的技巧與視覺藝術，是否令使用者方便操作與賞心悅目之外，更重要的是網站所使用的資訊技術，是否能帶給使用者耳目一新的感覺。因此，如何充分的運用資訊技術於數位化資料上，將是本子計畫追求的終極目標。以下就本子計畫在第一年計畫中所設定研發完成的工作目標，分別說明如下：

1. **高效率的書目檢索系統設計：**由子計畫一負責收集、組織、整理佛學書目與佛學論文資料後，提供給本子計畫作為書目檢索系統之資料庫。由子計畫二負責書目檢索網頁之美工設計，提供給本子計畫作為書目檢索系統之操作介面。本子計畫將同時採用全文檢索技術與欄位檢索技術，設計開發出一高效率的書目檢索軟體。並將此佛學書目檢索系統架設於全球網際網路之上，開放給一般民眾使用。

2. **高效率的佛教辭典檢索系統設計：**由子計畫一負責收集、組織、整理丁福保《佛學大辭典》資料後，提供給本子計畫作為佛教辭典檢索系統之資料庫。由子計畫二負責佛教辭典檢索網頁之美工設計，提供給本子計畫作為佛教辭典檢索系統之操作介面。本子計畫將採用欄位檢索技術，設計開發出一高效率的佛教辭典檢索軟體。並將此佛教辭典檢索系統架設於全球網際網路之上，開放給一般民眾使用。

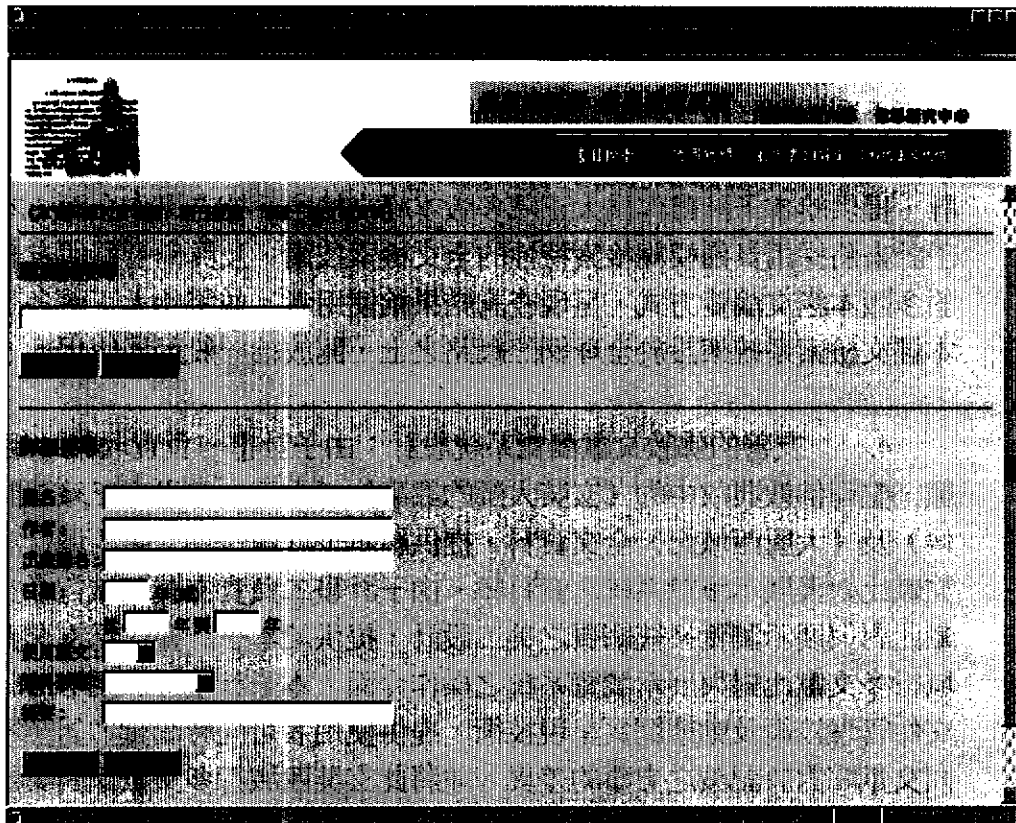
3. **智慧型的佛教經文檢索系統設計**：由子計畫一負責收集、組織、整理佛教經文中最具代表性的《大正新脩大藏經》全文資料與各家不同的翻譯版本，提供給本子計畫作為大正藏多版本經文檢索系統之資料庫。由子計畫二負責大正藏多版本經文檢索網頁之美工設計，提供給本子計畫作為大正藏多版本經文檢索系統之操作介面。本子計畫將採用具備模糊檢索特性的全文檢索技術，及具備智慧與友善的多版本經文瀏覽方式，呈現查詢結果給使用者。並將此大正藏多版本經文檢索系統架設於全球網際網路之上，開放給一般民眾使用。

4. **友善的佛教文獻瀏覽系統設計**：由子計畫一負責收集、組織、整理佛教典故中的《玄奘大師的故事》、《大唐大慈恩寺三藏法師傳》及《大唐西域記》全文資料，提供給本子計畫作為玄奘西域行多文獻比對與展示瀏覽系統之資料庫。由子計畫二負責玄奘西域行多文獻比對與展示瀏覽系統網頁之美工設計，提供給本子計畫作為玄奘西域行多文獻比對與展示瀏覽系統之操作介面。本子計畫將採用 XML 文件管理技術及動態網頁呈現技術，提供使用者可同時瀏覽並對照不同文獻間對於記載之史事的差異。並將此玄奘西域行多文獻比對與展示瀏覽系統架設於全球網際網路之上，開放給一般民眾使用。

(二)、完成之工作項目

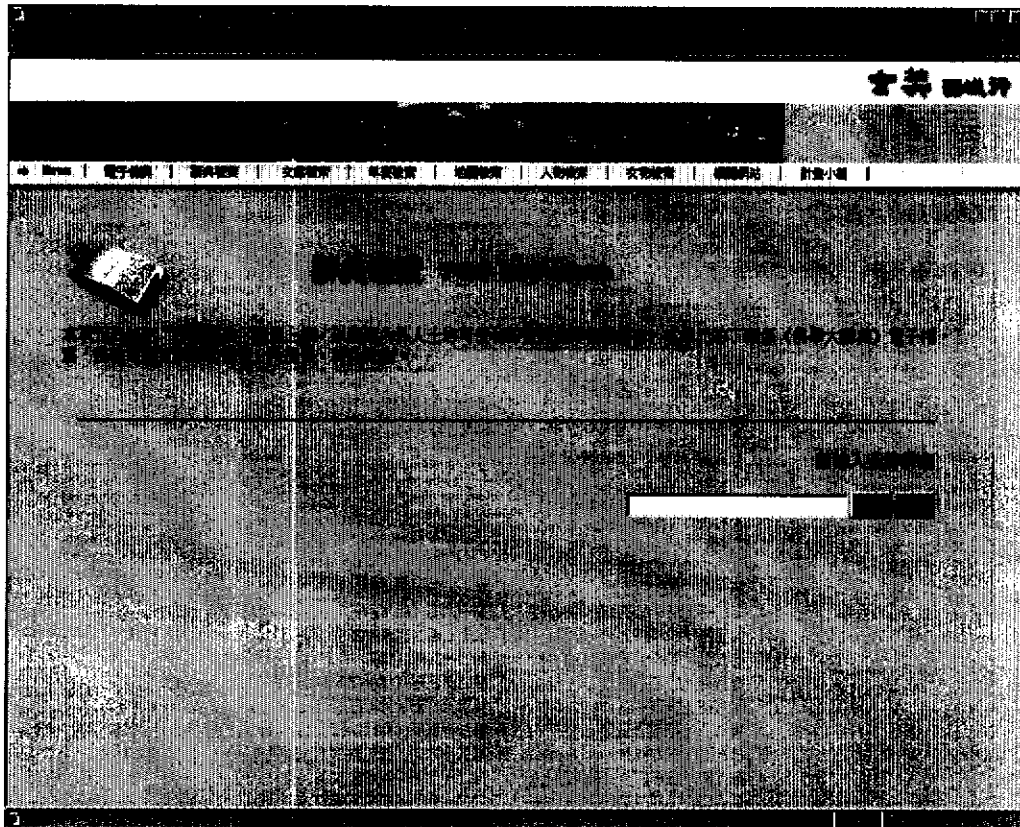
第一年我們完成的工作項目有「玄奘西域行數位博物館」的佛學書目檢索系統、佛教辭典檢索系統、大正藏多版本經文檢索系統及玄奘西域行多文獻比對與展示瀏覽系統之設計與開發。以下就本子計畫在第一年計畫中所完成的各項系統，分別說明如下：

1. **佛學書目檢索系統**：本系統主要乃蒐集佛學相關書目、期刊雜誌與博碩士論文等資料，提供一完善的查詢介面，供使用者查詢佛學方面的書目資料。在檢索技術的研發方面，我們除了設計全文關鍵字檢索技術之外；並且設計出欄位檢索技術，方便使用者可指定書目欄位作為檢索依據。此一系統的開發完成將帶給社會大眾享用資訊技術的便利性與即時性，使用者將可快速地、有效率地、方便地查詢佛學書目資料。



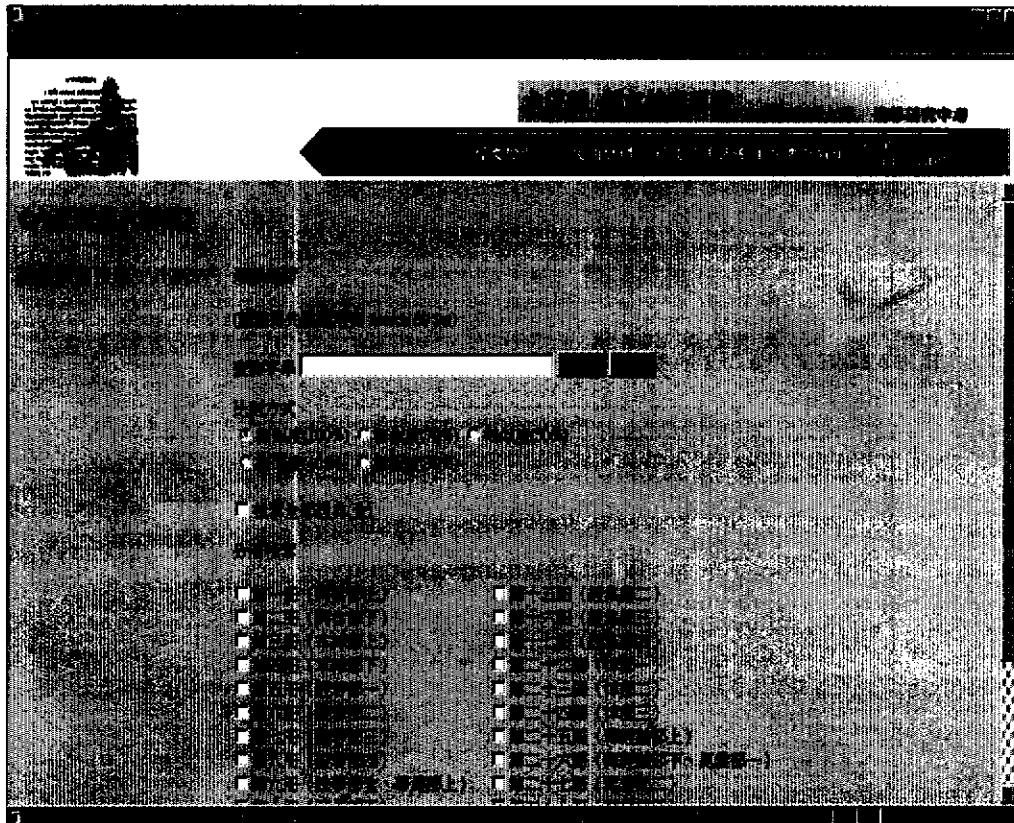
佛學書目檢索主畫面

2. 佛教辭典檢索系統：本系統主要乃蒐集丁福保《佛學大辭典》資料，將整個辭典電子檔經過必要之轉換程序，後加上檢索功能，提供使用者查詢佛學方面的專有辭彙。在檢索技術的研發方面，我們使用 XML 欄位檢索技術。此一系統的開發完成對佛學研究有興趣的社會大眾，是一不可或缺的工具。

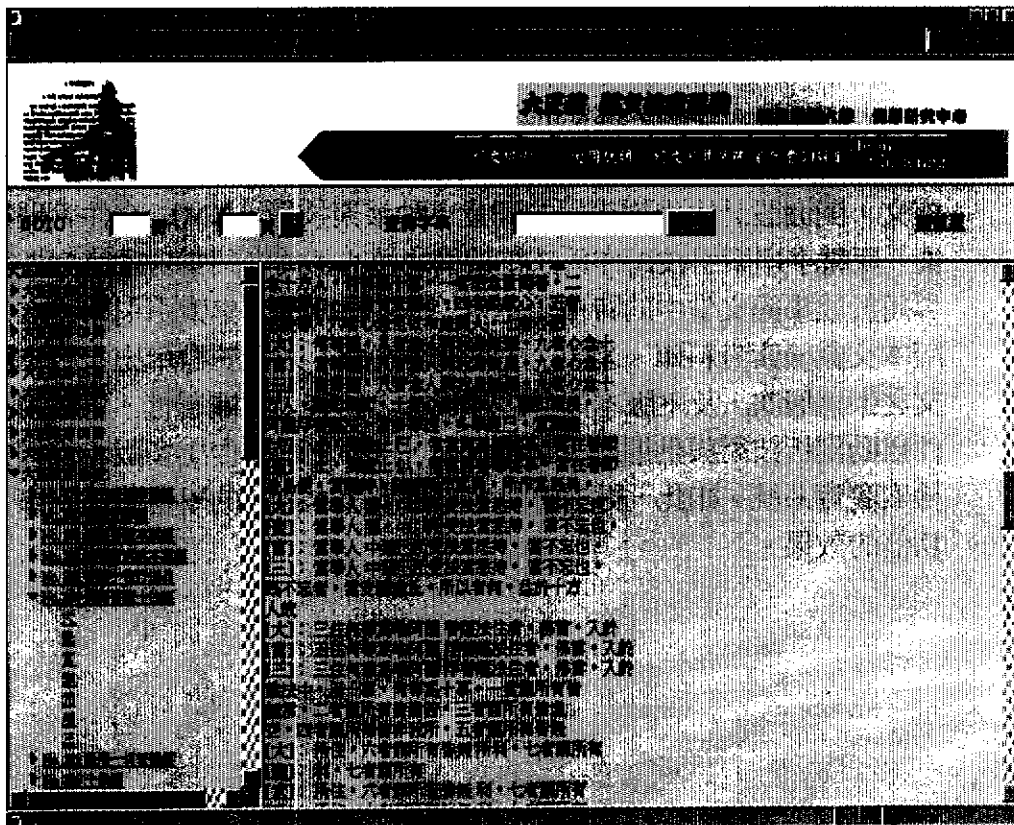


佛教辭典檢索主畫面

3. 大正藏多版本經文檢索系統：主要乃依據在不同朝代、不同翻譯者對於《大正新脩大藏經》經文的各翻譯版本為資料，整合各版本之差異性，提供彈性，實用性之瀏覽檢索方式予使用者。本系統針對數位化經文，開發設計出佛經專用的詞彙自動叢聚技術、XML 文件管理技術、反轉字串索引技術作為系統架構的核心。其中詞彙自動叢聚技術專用於收集經文當中的特徵值集合，提供給反轉字串索引引擎建立多版本文獻檢索資料庫。XML 文件管理技術，則用於標示出多版本經文之架構層次，及主版本與各翻譯版本之內容上的關聯性。再使用文件自動剖析技術與動態網頁呈現技術，自動化的製作出多版本佛經文檢索瀏覽系統。如此，對於研究大正藏多版本經文的學者將達到事半功倍的效果。



大正藏經文檢索主畫面



多本經文瀏覽主畫面

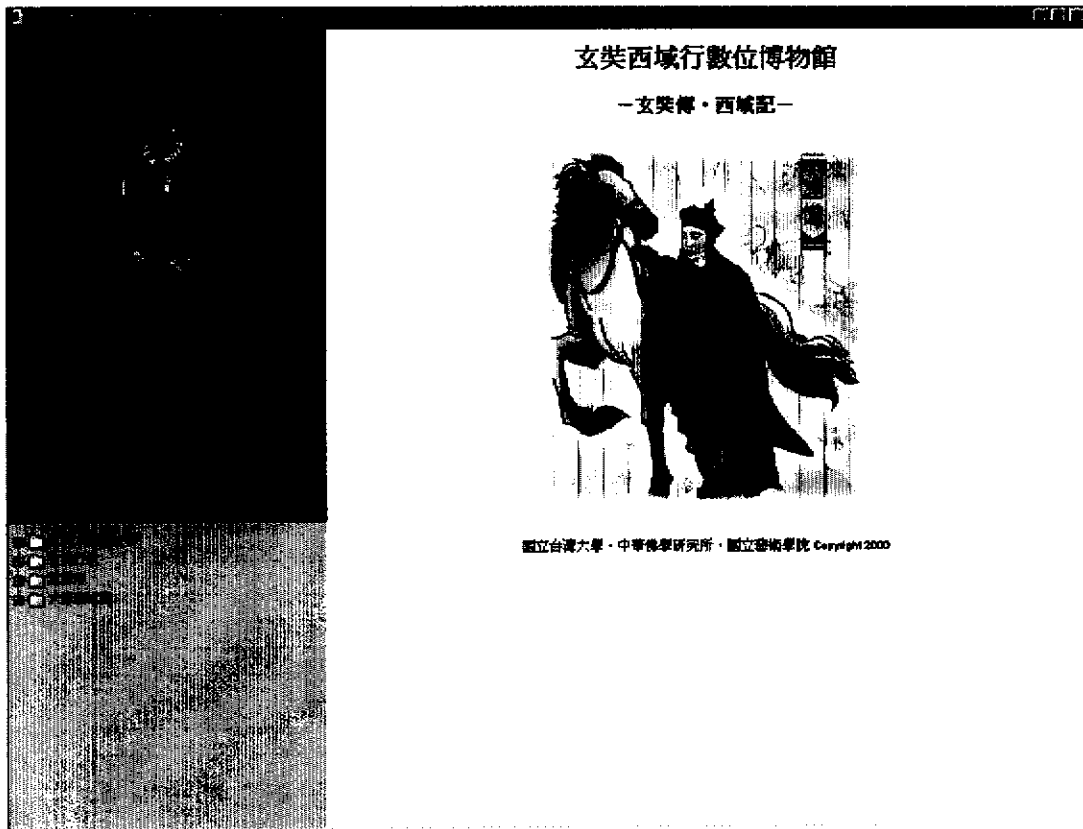
序號	詞彙	次數	類別
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16

最常查詢詞彙主畫面

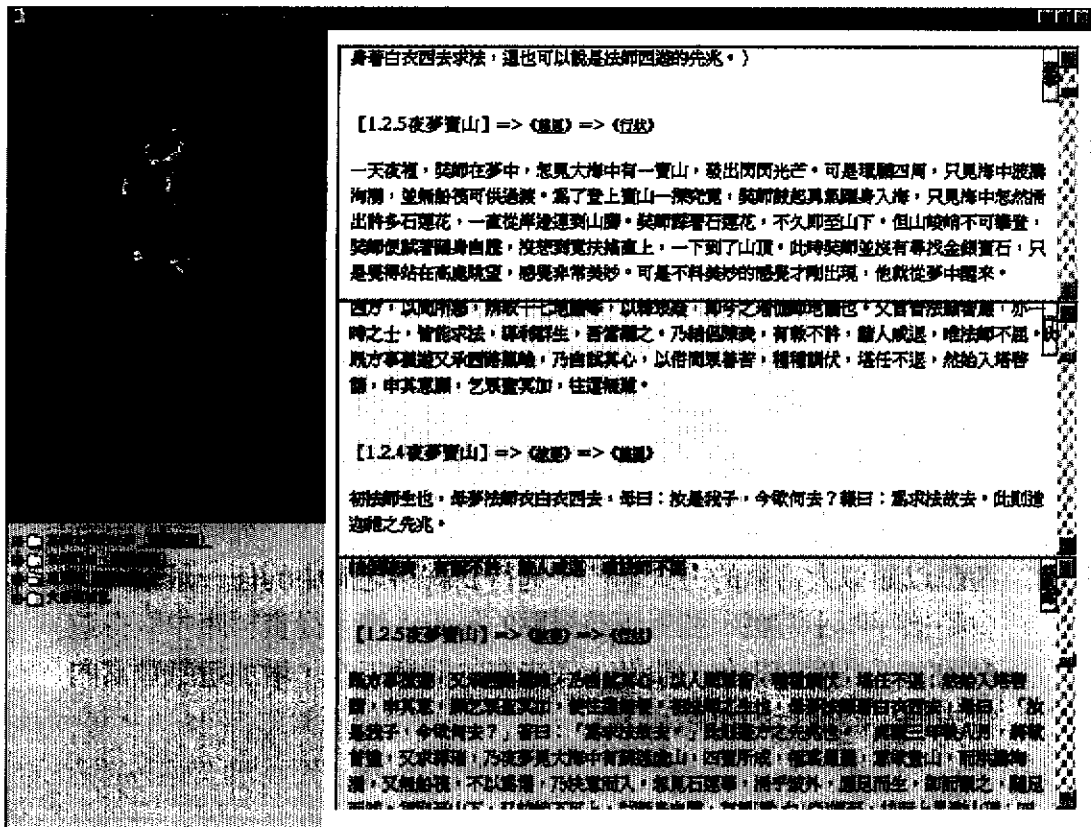
Cluster	詞彙
Cluster 1	...
Cluster 2	...
Cluster 3	...
Cluster 4	...
Cluster 5	...
Cluster 6	...
Cluster 7	...
Cluster 8	...
Cluster 9	...
Cluster 10	...

詞彙彙聚主畫面

4. 玄奘西域行多文獻比對與展示瀏覽系統：主要乃依據大唐時期玄奘西域行之典故為背景，蒐集《玄奘大師的故事》、《玄奘行狀》、《慈恩傳》及《大唐西域記》這四部佛教文獻，採用 XML 文件管理技術，將這每一部文獻建立出個別的樹狀目錄，並且整合這四部佛教文獻建立出彼此之間在內容與段落上的關聯性。再使用文件自動剖析技術與動態網頁呈現技術，自動化的製作出多文獻比對與展示瀏覽系統。如此，對於研究佛教文獻的學者將可專注於文獻的研究之上；而不用浪費時間於瑣碎的書本翻閱上。本系統的最大特色就是可同時呈現此四部佛教文獻相關段落的文章，使用者在透過不同文獻段落的閱覽與比對之後，對於研究佛教典故將會有事半功倍的感受與獲益良多的感觸。



玄奘西域行多本佛典瀏覽系統主畫面



文英西域行多本佛典瀏覽系統主畫面

5、「互動式學習」

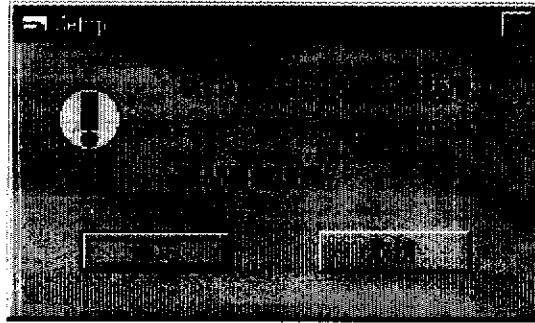
1.)「互動式學習」程式開發

為使使用者能更容易接受數位博物館的資訊，使數位博物館的「數位」二字真正能發揮其功能，而不只是將原本的經典文物轉為電子化，我們設計了互動式學習程式，讓使用者能以最輕鬆、最直接的方式利用數位媒體的聲光互動效果，達到最高的學習成效。為達到這樣的目的，我們使用 Java 設計本互動式學習遊戲，由於 Java Applet 在設計上、功能上都較 HTML 能有更多的彈性、空間，可輕鬆地整合聲音、動畫、圖片、3D Model、影像等等。我們的互動式遊戲在設計前也著眼於此，為充分發揮 Java 的強大功能。另一重點則為 Java 的跨平台特性，使無論在任何有連結網際網路的機器都可以使用我們的互動式遊戲。

2.)「互動式學習」程式的特色

本程式不同於一般網路上的 Java Applet 小遊戲程式，為容納絲路之旅的大量數位文物資訊，與整合多種不同遊戲內容，設計時使用 JDBC 技術，使用者在使用時無須等待下載龐大資料的時間，只需在

遊戲進行需要時，從 Microsoft Access 資料庫中取得即可。對於設定 ODBC 部分，我們也設計了一個簡單的程式，使使用者僅需按下紐，即可在電腦中自動完成資料庫的設定。



遊戲以類似大富翁的方式進行，進行到絲路之旅中的任一景點，均會隨機呼叫小遊戲，並對應該景點的風土人情，從資料庫中取得遊戲內容，高度模組化的設計，使未來的開發工作，增加遊戲豐富內容方面，能進行的更順利。

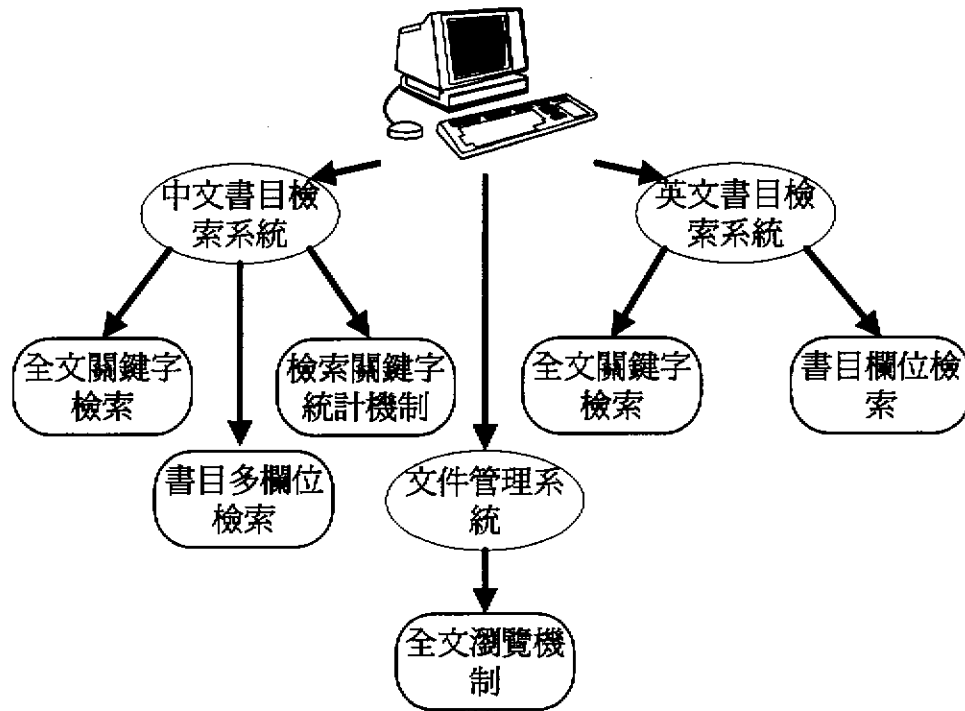
3.) 完善的使用者介面

為使本遊戲能達到互動學習的目標，本遊戲設定對象為不論男女老幼的一般大眾，遊戲進行前精采的 Flash 片段，更吸引使用者進行學習，本遊戲也具有詳盡的線上說明文件，使使用者能在最短的時間內上手。本遊戲的使用介面也儘量做到人性化設計，使用者甚至僅憑直覺，即可順利的享受互動式學習遊戲。

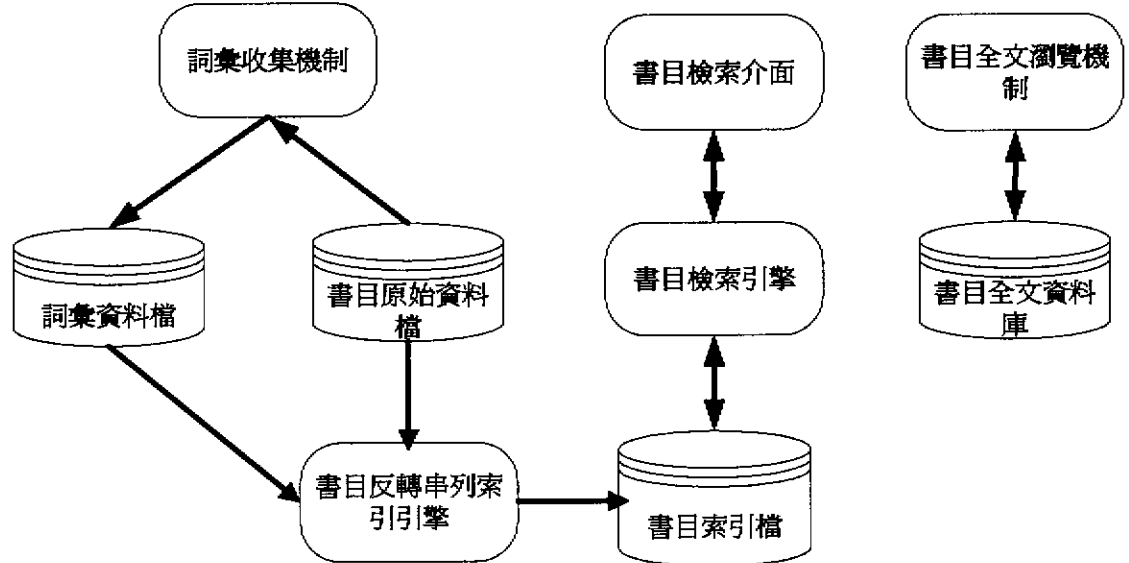
(三)、系統架構與檔案結構說明

茲將第一年所設計開發出來的系統，針對其系統架構與檔案結構，分別說明如下：

1. 佛學書目檢索系統： 系統架構圖



系統維護流程圖

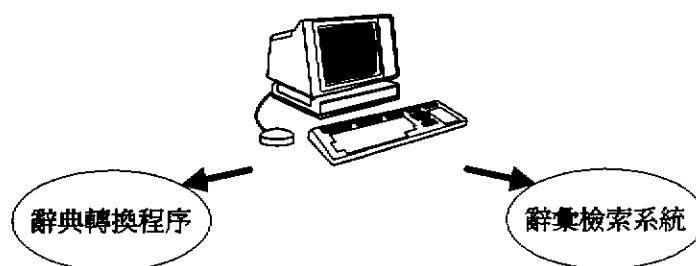


檔案結構表

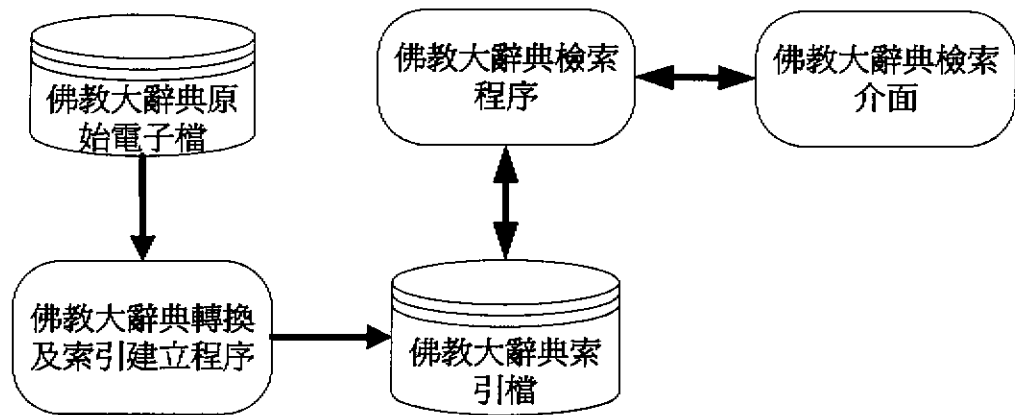
no	檔 名	內 容	註
1	Cserver	中文書目檢索引擎	
2	Eserver	英文書目檢索引擎	
3	PhraseIndex	辭彙索引集合	
4	PhraseInv	辭彙之反轉串列索引檔	
5	RecIndex	書目記錄邏輯磁碟偏移位置對照檔	
6	TotalFile	書目記錄全文資料檔	
7	TotalIndex	書目記錄全文反轉串列索引檔	
8	TotalInv	書目記錄全文索邏輯偏移位置對照檔	
9	Fields	書目欄位名稱對照檔	
10	Severer_nwlin	書目查詢結果呈現執行檔	
11	ChineseIndex.html	中文書目檢索首頁	
12	ChineseFind.html	中文書目檢索辭彙統計表首頁	
13	ChineseInfor.html	中文書目資料提供與建議首頁	
14	ChineseUse.html	中文書目檢索說明首頁	
15	English.html	英文書目檢索首頁	
16	Ereadme.html	英文書目檢索說明首頁	

2. 佛教辭典檢索系統：

系統架構圖



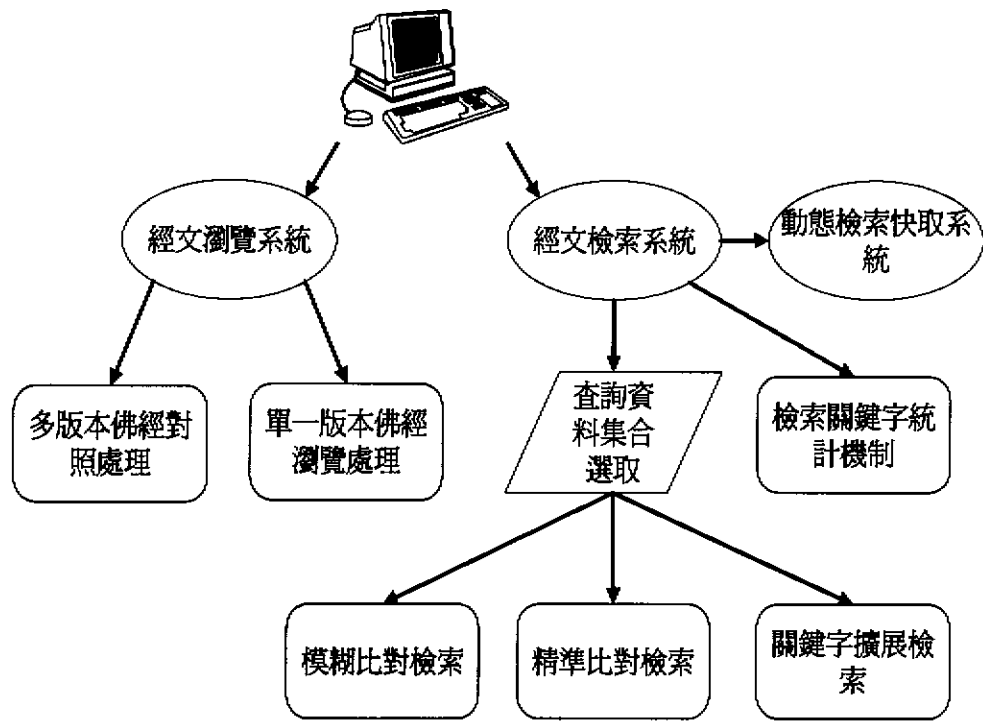
系統維護流程圖



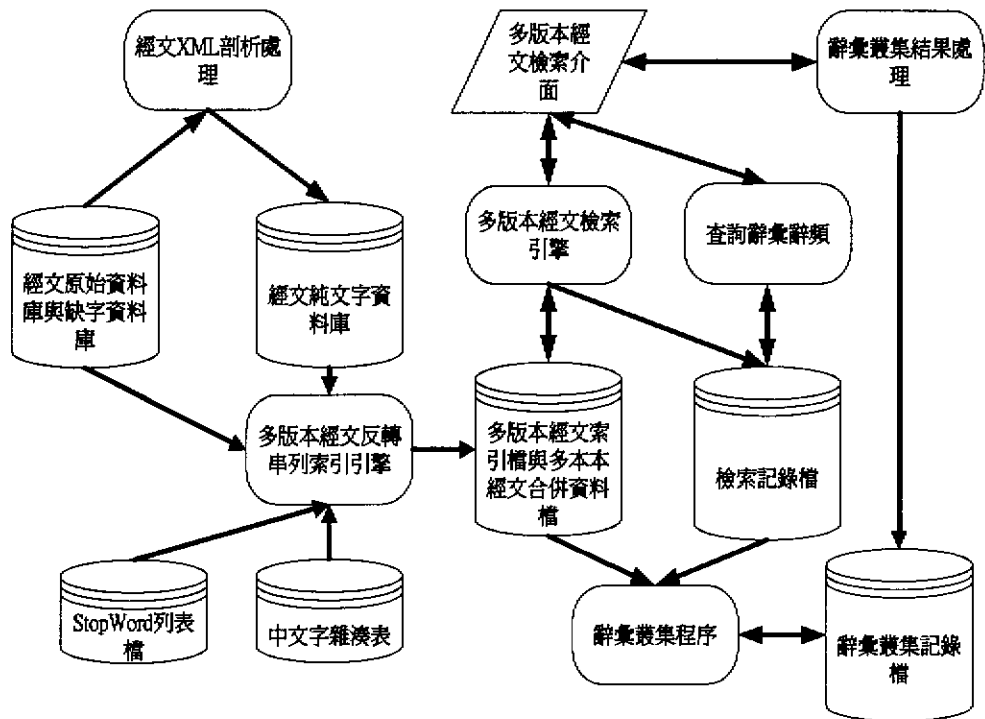
檔案結構表

no	檔 名	內 容	註
1	Dict.txt	丁福保《佛學大辭典》全文電子檔	
2	CreateDB.java	辭典轉換原始檔	
3	CreateDB.class	辭典轉換執行檔	
4	Ccbsdb.*	辭典索引檔	由 CreateDB.class 產生
5	Search.jsp	辭典索引引擎	
6	Db-dict.html	佛教辭典檢索首頁	

3. 大正藏多版本經文檢索系統：
系統架構圖



系統維護流程圖

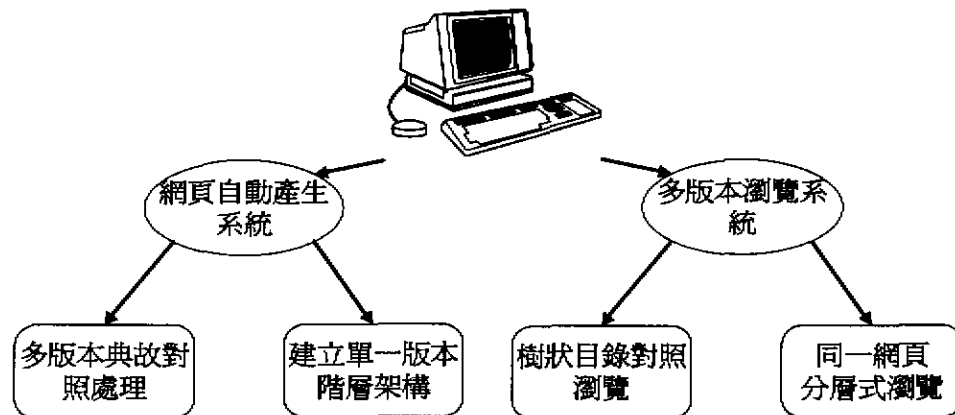


檔案結構表

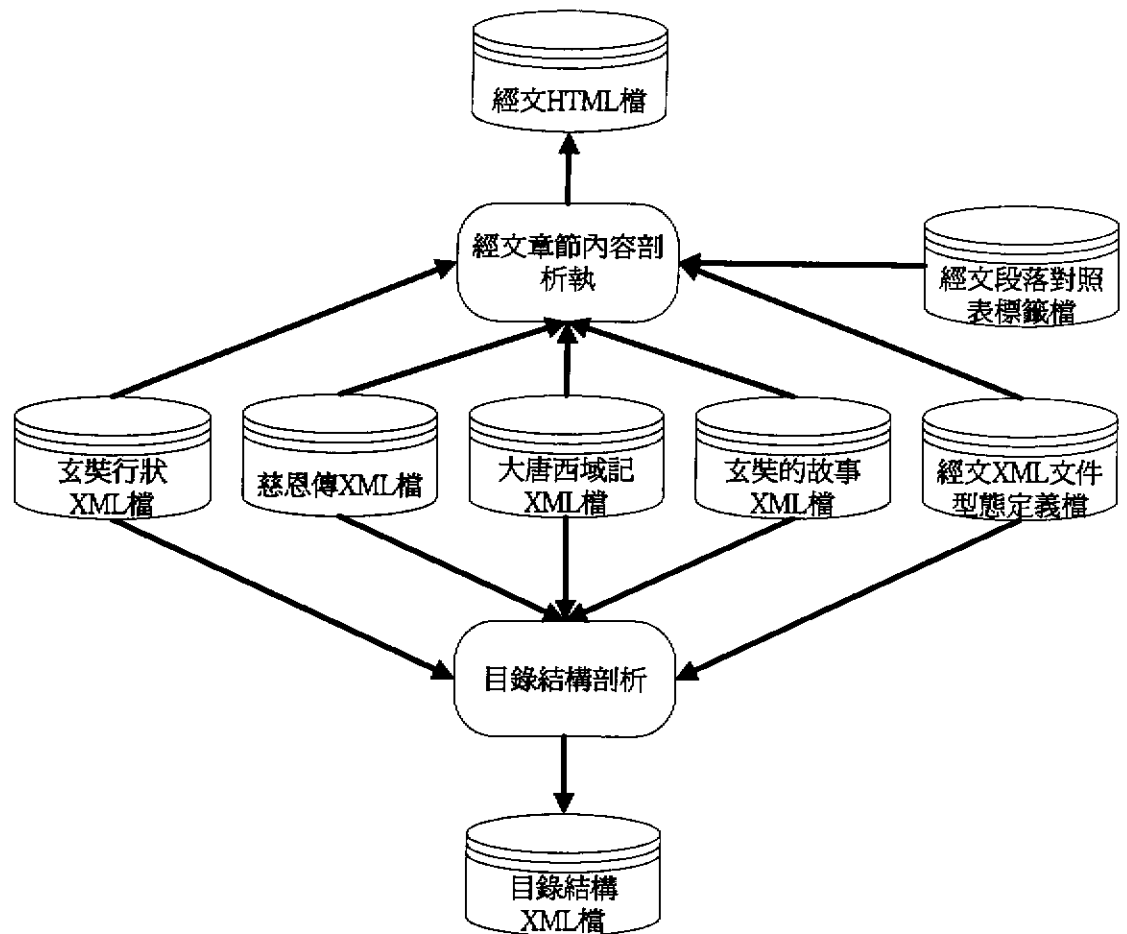
no	檔名	內容	註
1	T21/*.xml	大正藏經文 XML 檔	
2	T21/*.ent	大正藏經文缺字連結檔	

3	Parser.pl	多版本經文前置處理程式	輸入 T21/*.xml 及 T21/*.ent
4	Tresult/*main	多版本經文資料檔	由 Parser.pl 產生 經過去除標籤處理
5	WordToNum	中文字雜湊表	用於反轉串列之建立
6	Main.pl	多版本經文索引建立程式	輸入 Tresult/*main
7	InvertedFile	反轉串列檔	經由 main.pl 產生
8	RecIndex	經文記錄層次索引表	合併後之搜尋資料
9	FileIndex	經文檔案層次索引表	合併後之檔案
10	Brochure	經文分類層次索引表	
11	TotalFile	多版本經文資料	經過處理與合併 經由 main.pl 產生
12	SutraServ	多版本經文檢索引擎	
13	CGI/list.cgi	顯示查詢辭彙頻率	
14	Miclass/class.cgi	瀏覽辭彙叢聚結果之程式	
15	Index.html	大正藏經文檢索首頁	
16	Browse/cbeta.html	多版本瀏覽與查詢首頁	
17	Find.html	最長查詢辭彙首頁	
18	Term.html	顯示辭彙叢集首頁	

4. 玄奘西域行多文獻瀏覽系統：
系統架構圖



系統維護流程圖



檔案結構表

no	檔名	內容	註
1	2052.xml	玄奘行狀 XML 檔	
2	2053.xml	慈恩傳 XML 檔	
3	2087.xml	大唐西域記 XML 檔	
4	9999.xml	玄奘的故事 XML 檔	
5	Xml2dir.bat	目錄結構剖析執行檔	
6	X*.xml	目錄結構 XML 檔	由 xml2dir.bat 產生
7	Cbeta.dtd	典故 XML 文件型態定義檔	
8	Html.bat	章節內容剖析執行檔	
9	2052*.html	玄奘行狀 HTML 檔	由 Html.bat 產生
10	2053*.html	慈恩傳 HTML 檔	由 Html.bat 產生
11	2087*.html	大唐西域記 HTML 檔	由 Html.bat 產生
12	9999*.html	玄奘的故事 HTML 檔	由 Html.bat 產生
13	XSABRlink.xml	經文段落對照表標籤檔	

14	Menu.js	動態展示目錄結構執行檔	
15	Tang-index.html	玄奘西域行首頁	

二、子計畫三成果檢討

1. 絲路之旅(中亞、印度部分)內容待增加

目前互動式遊戲的內容中，遊戲進行的地圖僅達中國境內，絲路之旅中極重要的中亞、印度部分仍未加入遊戲中，而如何使原來的 Java 程式能夠更新龐大的地圖模組，並加入這一部分的新內容，是我們再第二期計畫中的重點。

2. 網路功能未竟完全

為希望使用者能完全達到理想中，「僅需連上網路，任何人都可以輕鬆的使用互動式學習」目標，我們現在的程式仍未達到這樣的理想，使用者必須由網路上下載我們的遊戲文物資訊資料庫，再瀏覽遊戲網頁，遊戲方可進行。我們希望開發新的互動式遊戲技術，使我們的互動式遊戲能連接網路上的 Microsoft SQL 資料庫，我們將採用的策略為於伺服器端架設 Microsoft IIS 站台，其上架構 Microsoft SQL 資料庫，內容包含原 Microsoft Access 遊戲文物資訊資料庫的內容。並改進我們的 Java 互動式遊戲程式，使其能夠與 IIS 上的 SQL 資料庫溝通，如此一來，我們在遊戲前即可省去下載資料庫、執行 Setup 程式的麻煩，直接進行遊戲。另外原遊戲是架設於 Microsoft PWS(Personal Web Server)上，一旦使用者人數成長，可能會無法負荷的問題，也可於此一併解決，這在第二期工作中，為後期進行的重點。

互動式遊戲進行畫面

