

# 以 QTI 為標準引入 IRT 之試題編輯工具

廖元偉 趙銘

逢甲大學資訊工程研究所

wway@mail.mmlab.iecs.fcu.edu.tw

mingchao@fcu.edu.tw

## 摘要

本文旨在開發一套數位化適性題庫編輯模組 (Adaptive Itembank Editing Module, AIEM)。AIEM 的設計主要是以模組化的概念建置於開放原始碼的學習管理平台—MOODLE。AIEM 主要針對適性化和試題互通性為主，發展成為一套支援 IMS QTI (Question and Test Interoperability) 試題交換標準格式的匯入的編輯模組，使題庫的開發方面更為便利。AIEM 模組的架構主要分為五個功能元件、三個樣版、及資料庫，分別為「題庫編輯元件」、「題庫瀏覽元件」、「試題編輯元件」、「試題瀏覽元件」、「分析轉換元件」、「題庫資訊樣版」、「試題資訊樣版」、及「詮釋資料樣版」。

**關鍵詞：**題庫編輯、IMS QTI、IRT、MOODLE

## Abstract

We implement an Adaptive Itembank Editing Module (AIEM), and integrate AIEM with open source Learning Management System MOODLE (Modular Object Oriented Dynamic Learning Environment). AIEM is focus on adaptive test and interchange item that is. based on IMS QTI. We try to develop itembank editing module, which is more convenient to develop itembank. The architecture of AIEM includes five functional components, three templates, and one database that are Itembank editing component, Itembank browsing component, Item editing component, Item browsing component, Analyses and transform component, Itembank template, Item template, and Metadata template.

**Keywords:** itembank, IMS QTI, IRT, MOODLE

## 1. 前言

適性化測驗將成為未來電腦化測驗的主流，然而適性化測驗所仰賴的是需要龐大的題庫資源。本文主要是建置一套能夠符合 QTI (Question and Test Interoperability) 標準且引入試題反應理論 (Item Response Theory, IRT) 觀念的數位化適性題庫編輯模組 (Adaptive Itembank Editing Module, AIEM)。AIEM 主要是要達到以下目標：

1. 與 MOODLE (Modular Object-Oriented

Dynamic Learning Environment) 學習管理平台結合的 Web-base 的數位化適性試題編輯模組。

2. 以 QTI 標準的詮釋資料做為基礎結合試題反應理論來管理試題庫。
3. 轉換試題為符合 QTI 標準格式的 XML 檔案，並使其匯出達到試題再利用。

測驗和評量一直是教學過程中不可或缺的一環，透過測驗的結果不但可以反應出學生對學習內容了解的程度，提供學生自行參考或者讓教師掌握學生的學習狀況，更可以利用測驗的結果達到能力的判定，例如：托福考試。

電腦化測驗和傳統紙筆測驗有著許多不同特性，然而使用電腦化測驗就字面上來說符合環保的意識便是其中之一的優點，然而不止於此，電腦化測驗也可節省測驗投入的成本、提供立即回饋訊息、及提升工作效率等。電腦化測驗可在教育測驗的領域裡，改善許多傳統測驗較難完成的工作。

在電腦化測驗中，適性測驗也是令人關注的研究，在傳統的測驗中，相同長度與相同試題的試卷，給予不同能力的學生進行測試，不但浪費時間，也容易使能力較差的學生感到挫折。適性化測驗能改善傳統測驗所造成的不便，但由於適性化測驗需要針對不同的個人給予不同的試題，在傳統的紙筆測驗模式中，實施適性化測驗有一定的困難所在，較難普遍的使用適性化測驗[3]。在網際網路與個人電腦發達的現在，適性化測驗的發展逐漸明朗，教學者可以建立題庫讓學習者透過網際網路進行線上測驗；而系統將會針對不同學習者的能力進行選題，最終達到能力值的判定。

隨著數位學習的發展，學習管理系統發展也漸趨於成熟，因受限於版權問題，AIEM 主要使用開放原始碼的學習管理平台—MOODLE，並參照 IMS QTI 標準和 IRT 觀念。

## 2. 文獻探討

此小節主要探討開發 AIEM 時所相關的文獻，分為 IMS QTI、IRT、及題庫建置流程三個部分。

### 2.1 IMS QTI

IMS (Instructional Management System) 全名為全球學習聯合機構 (IMS Global Learning Consortium) 是一個非營利性的組織，主要在發展

和訂定教育與學習科技相關的各項規格，已有許多制定的規格成為公認的標準，其所發表的 QTI，主要描述考題、試卷、及評量結果之規格，目前最新版本是 2.0 於 2005 年一月所發佈，它是以可延伸式標籤語言 (eXtensible Markup Language, XML) 來定義相關規格內容，此份 QTI 規格制定的主要目的在於用以解決平台之間測驗資訊交換、管理、及搜尋之問題，並達到不同平台間評量資源能彼此交流和共享[11]。

QTI 規格主要分為下列兩個部分：

1. ASI Model (Assessments Sections Items Model)：定義試題的類型、試題組成方式、選題的方式、及成績計算方式等。
2. Results Reporting：定義如何呈現學習者的測驗結果。

QTI 規格書提供了高度的延伸性 (extensible) 和客製化 (customizable)，所以一個特定的系統可以容易快速地適應以及應用這種規格，相反地，也可以根據自己的需求而增加所必需的相關資訊。

## 2.2 試題反應理論

測驗理論 (test theory) 是一種解釋測驗資料間實證關係 (empirical relationships) 的有系統的理論學說，其發展測驗理論學者通常把它劃分成二大學派：一為古典測驗理論 (classical test theory)，主要是以真實分數模式 (true score model) 為骨幹；另一為現代測驗理論 (modern test theory)，主要是以 IRT 為架構，試題反應理論始於 1990 年 Lord 發表第一本以「試題反應理論」為名的專書後於是正式正名，測驗學者逐漸以試題反應理論為現代測驗理論的代表。這兩派理論目前並行流通於測驗學界，但試題反應理論卻有後來居上，逐漸凌駕古典測驗理論之上，甚至進而取代之勢[1]。

反應模式是以數學模式來表示受試者和能力之間的關係。模式的分類若依據試題特徵曲線的呈現方式，可以分為參數型試題反應模式與無參數型試題反應模式；若以測驗資料的性質而定，也就是以測驗計分方法的不同來分類，可分為二元計分、多元計分、及連續資料等模式。參數型試題反應模式較常用的是對數模式，有單參數 (公式 1)、雙參數 (公式 2)、及三參數 (公式 3) 模式，每一種模式是依其採用的試題參數的數目多寡來命名，適用於二元化的反應資料：

$$p_i(\theta) = \frac{e^{D(\theta-b_i)}}{1 + e^{D(\theta-b_i)}} \dots\dots\dots(\text{公式 1})$$

$$p_i(\theta) = \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}} \dots\dots\dots(\text{公式 2})$$

$$p_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta-b_i)}}{1 + e^{Da_i(\theta-b_i)}} \dots\dots\dots(\text{公式 3})$$

其中， $i=1, 2, \dots, n$ ， $D=1.7$

$\theta$ ：受試者能力值

$p_i(\theta)$ ：表示具能力  $\theta$  的受試者答對第  $i$  個試題的機率

$a$ ：表示第  $i$  個試題之鑑別力參數

$b$ ：表示第  $i$  個試題之難度參數

$c$ ：表示第  $i$  個試題之猜測參數

常數  $D$  是為了讓試題反應模式接近於常態肩形模式 (normal ogive model)。單參數對數模式被視為三參數對數模式的特例，因其假設猜測參數  $c$  等於 0 及鑑別力參數  $a$  等於 1 時，兩種模式便沒有什麼差別了[2]。

## 2.3 題庫建置流程

數位化適性題庫編輯模組主要是提供一個適性題庫的發展環境，並達到試題的共享性，在傳統適性題庫的建置過程中，試題參數的收集主要以傳統紙筆測驗的方式為主，在[9]的討論之中，證明也可以藉由網際網路的方式來分析試題的各項參數，因此本題庫編輯模組，捨棄傳統試題參數的收集方法，在設計時使用網際網路為收集試題參數數據為導向，重新規劃試題建置流程。圖 1 為數位化適性題庫編輯模組題庫建置流程，由上而下，步驟一到步驟四主要是在於建立題庫的資訊，描述題庫的建置目的，且藉由雙向細目表的建置，預先規劃題庫的內容。若以傳統的方式收集數據，在預試前並不需要先行建置數位化的題庫，而是在預試之後修正參數再建置數位化的題庫，因使在本系統將在試題編製時提供參數的輸入，用以建立適性題庫，並藉由網際網路的途徑來進行預試。

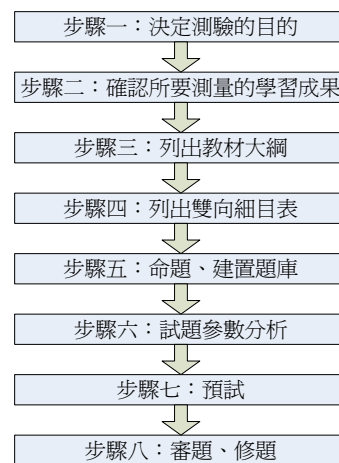


圖 1 數位化適性題庫編輯模組題庫建置流程

在編輯題庫時首先要決定測驗目的，測驗的編製計畫會因測驗目的的不同，而有不同的考量，因此在進行測驗的規劃時，必需清楚測驗的目標為何，一般來說測驗的類型可分為四類，如表 1[7]。

表 1 四類成就測驗的特徵[7]

測驗的類型	測驗的功能	抽樣的考慮	題目的特徵
安置測驗	1. 測量學習前所需之知識和能力； 2. 決定學習之前，已達到課程目標之程度。	1. 包括各種學習前所需之起點行為； 2. 選擇課程目標之代表性樣本。	1. 通常題目之難度低，是效標參照的； 2. 題目難度之分數範圍較廣，是常模參照的。
形成性測驗	為學生及教師提供學習進步情形之回饋。	如果可能則包括所有單元目標或最必要的目標。	題目配合單元目標的難度，是效標參照的。
診斷測驗	診斷學習困難之原因。	以共同的學習錯誤為依據。	通常，項目之難度低，用來指出特殊學習錯誤之原因。
總結測驗	在教學結束時，用來評定等第或證明以熟練教材。	選擇課程目標的代表性樣本。	通常題目難度之分數範圍廣，是常模參照的。

接著必需確認所要測量的學習成果，一個測驗所測量的學習結果，應能忠實地反應教學的目標。基於此，編製測驗時首先要選定測驗所預測量的教學目標，其次要使陳述目標的方式適合於測驗之編製。在分析並陳述教學目標時，最適當的參考資料是布魯姆等人所著的「教育目標的分類」。他將教育目標分為認知的(cognitive)、情意的(affective)、及動作技能(psychomotor)三大領域。其中認知領域主要是關於智性的學習結果，情意領域是有關興趣和態度的目標，動作技能顧名思義就是有關動作技能的目標。在認知的領域當中又可分為六個主要的範圍，分別是知識、理解、應用、分析、綜合、及評鑑，由簡單趨向複雜。

雙向細目表的目的是在於發現學習結果和課程內容的關係，並訂出各不同領域之相對比重，以確保測驗能測量到我們所預期之學習結果和教材的代表性樣本，再依據雙向細目表來編製測驗試題，配合特定的學習結果，並考慮到測驗的難度。

### 3. 適性題庫編輯模組分析

本模組藉由元件化的概念，試著設計一個數位化適性題庫編輯模組，提供給需要建置適性題庫的人員一個便利的工具，並結合試題的標準 IMS QTI 達到試題的再利用。

#### 3.1 系統執行環境

圖 2 為 MOODLE 的系統模組[6]，在 MOODLE 平台裡主要有三大功能：網站管理、學習管理、及課程設計。然而每項功能都是由許多模組所組成，例如：在課程設計當中具有作業模組、課程模組、及測驗模組等，本模組主要是配合課程設計功能裡的測驗模組進行設計。

Moodle

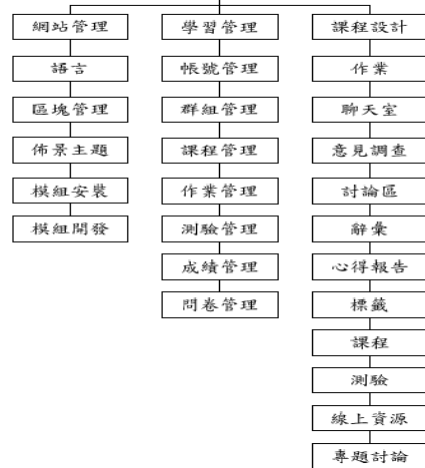


圖 2 MOODLE 系統模組[6]

圖 3 為系統環境示意圖，伺服器是指一般架設網頁的電腦，MOODLE 是指架設在伺服器之上的學習管理平台，透過伺服器發佈於網際網路。教師根據教案的教學目標來設計試題，然後再透過網際網路與 MOODLE 裡的課程設計功能裡的測驗模組進行編輯試題的工作，編輯結束後再利用 MOODLE 和伺服器進行溝通將資料儲存在伺服器裡的資料庫當中。管理員(系統維護員)透過網際網路登入到 MOODLE 學習管理平台之中，對於系統或題庫進行管理的動作。

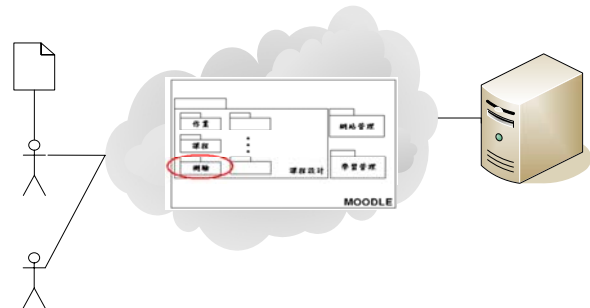


圖 3 系統環境示意圖

根據一些關於題庫編輯系統的文獻[4][5][8][10]。歸納題庫編輯系統的功能必需達到以下三點功能：

1. 試題的編輯：透過題庫編輯系統或工具，以文字、圖形、聲音、動畫、及影片等方式來編輯試題，並提供試題參數的輸入。
2. 試題再利用：提供匯出匯入檔案的功能，轉換試題格式，達到試題再利用。
3. 審題和修題：當完成預試，搜集到受試者的反應之後，可以藉由系統進行試題數據的分析，並修改試題題目或是試題參數。

### 3.2 模組架構

圖 4 為 AIEM 的模組架構圖，根據主要分為五個功能元件、三個樣版、及資料庫，分別為「題庫編輯元件」、「題庫瀏覽元件」、「試題編輯元件」、「試題瀏覽元件」、「分析轉換元件」、「題庫資訊樣版」、「試題資訊樣版」、「試題資訊樣版」、及「詮釋資料樣版」。

以下針對各個模組與樣版做進一步的描述。

1. 題庫編輯元件：此元件主要負責題庫的新增、修改、及刪除，包含了題庫的使用者介面，在使用者點選選擇後針對所選進行動作，在新增題庫時會先將題庫樣版匯入，提供使用者輸入題庫的相關資訊，然後載入詮釋資料樣版，形成介面供使用者輸入，接著再將資料匯入資料庫儲存。在修改時會先呼叫題庫瀏覽元件進行題庫的列表，然後根據使用者的選擇將題庫的資訊從資料庫中提出再匯入題庫樣版，供使用者修改，在修改完成後更動資料庫中的資料。
2. 題庫瀏覽元件：此元件主要的功能就是針對題庫進行列表和基本的關鍵字搜尋題庫，在使用者選擇瀏覽題庫後，可選擇列表或者是搜尋，在選擇列表後，題庫瀏覽元件會將資料庫中的資料提出並依照格式進行列表，若選擇搜尋，則根據題庫樣版詮釋資料的資訊，提供關鍵字的搜尋並回報搜尋的結果。
3. 試題編輯元件：此元件主要的功能為新增試題、修改試題、及刪除試題並包含試題編輯的介面，此模組一開始會載入試題樣版，形成介面供使用者輸入試題，在輸入完試題後再載入詮釋資料樣版，形成詮釋資料輸入的介面，然後再匯入至資料庫儲存。若使用者選擇修改試題，則會先行執行試題瀏覽元件，進行列表或搜尋，在針對使用者的選擇，從資料庫中提取資訊，與樣版結合形成介面提供給使用者修改。
4. 試題瀏覽元件：此元件主要進行試題的列表和搜尋，選擇列表後則直接從資料庫中提取資料，根據試題樣版資訊將資料列出。若使用者選擇搜尋，則會提供關鍵字的搜尋，並根據試題樣版和詮釋資料的樣版的資訊進行資料庫搜尋，在提供列表給使用者。
5. 分析轉換元件：此元件主要負責題庫的匯出與匯入。在此模組中，將會先行執行題庫瀏覽元件，在使用者選擇題庫再從資料庫中提取，並根據 QTI 的格式進行轉換。若選擇匯入，則會提供檔案上傳的介面，在檔案上傳後進行分析，將資料轉換後匯入資料庫。
6. 題庫資訊樣版：此樣版主要是提供題庫的相關資訊，例如：題庫編號、題庫名稱、及題庫描述等。
7. 試題資訊樣版：此樣版主要具有兩個部份的資料，分別為試題的類型和及其對應的試題資訊，例如：選擇題、多選題、試題層次、試題

反應模式、難度參數、鑑別度參數、及猜測度參數等。

8. 詮釋資料樣版：此樣版主要根據 QTI 的詮釋資料規定製作。

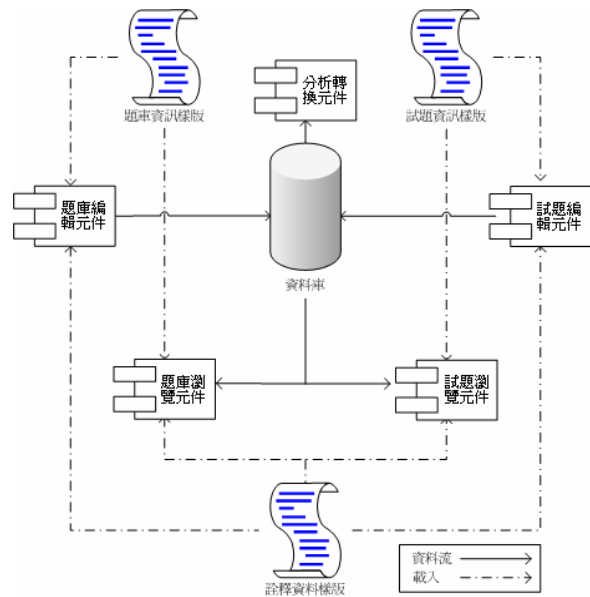


圖 4 AIEM 模組架構圖

### 4. 系統實作

本節主要是描述題庫的相關實作，包括題庫建置和試題產生兩個部分，在 UML 的圖形中，順序圖主要是描述訊息的傳送、執行時間、及模組或元件的關係，因此接下來的小節當中，將以順序圖和系統實際執行畫面來描述數位化試題編輯模組。

題庫的操作主要是透過題庫編輯元件來進行控制，在程式的部分，預先建立題庫的資訊樣版和詮釋資料樣版，圖 5 為新增題庫的順序圖，藉由圖 5 可以看出，在題庫編輯元件開始進行時，首先呼叫題庫資訊樣版之後再行載入，並形成介面提供輸入相關的設定，接著會在進行呼叫，所呼叫的變數已更改為詮釋資料的樣版，並載入詮釋資料的樣版。

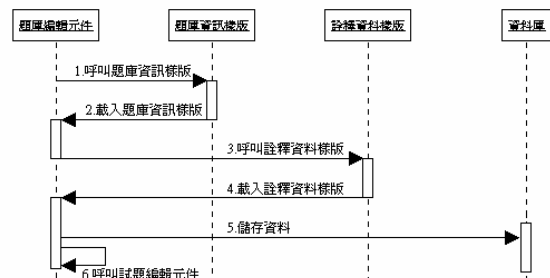


圖 5 新增題庫順序圖

圖 6 為試題資訊的輸入頁面，頁面中主要提供題庫編號、題庫名稱、測驗目標、題庫簡介、及雙向細目表的相關資訊。在題庫編號部分將由系統依據資料庫的資料順序產生編號；題庫名稱提供一個

針對題庫簡單的命名，例如：作業系統期中考；測驗目標提供對測驗的整體目標作一簡單的描述，例如：此份試題為作業系統的種類和功能，藉由測驗希望能使學生了解到自己對作業系統歷史的了解程度；雙向細目表的欄位上，可以透過編輯工具中的功能「表格」來達到繪製雙向細目表。

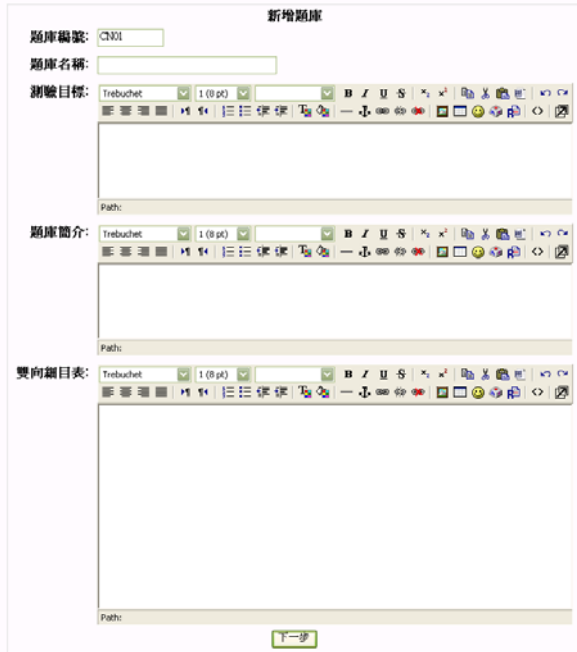


圖 6 題庫資訊輸入頁面

圖 7 為詮釋資料樣版載入後的執行結果，主要提供八大項的標籤輸入，IMS QTI 的詮釋資料主要是參照 IEEE LOM，取其中八大項的標籤。

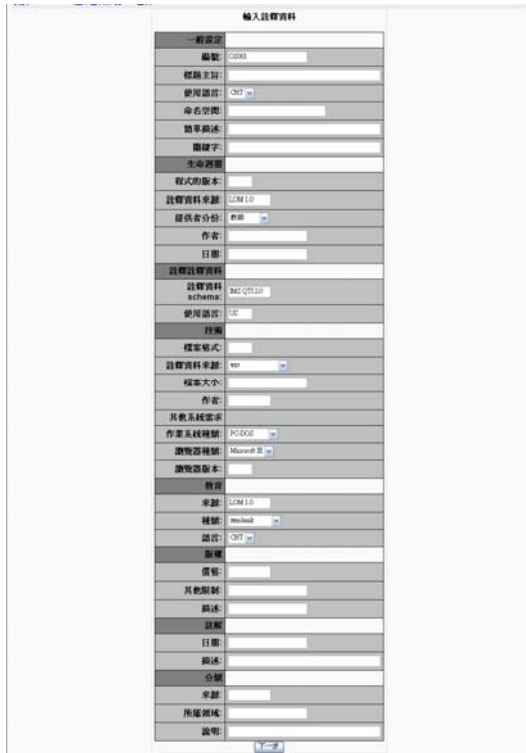


圖 7 詮釋資料輸入頁面

試題的產生主要是利用試題編輯元件來實踐，如圖 8 所示，在新增試題時必需先行選擇試題類型如圖 9，依據所選擇的試題類型動態的呼叫並載入試題資訊樣版，如圖 10 為是非題樣版執行後的畫面，然在再呼叫並載入詮釋資料樣版針對試題填入相關的詮釋資料，詳細的詮釋資料樣版參照圖 7。

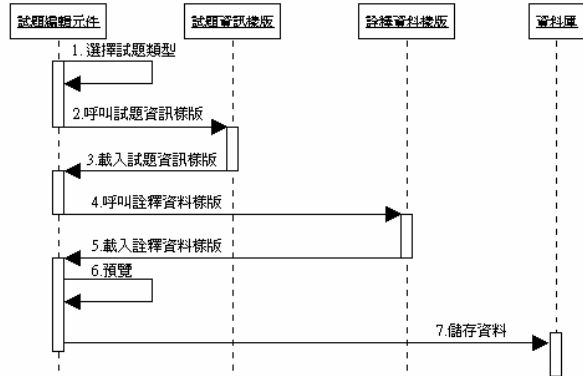


圖 8 新增試題順序圖



圖 9 選擇試題類型



圖 10 是非題試題資訊

表 2 為各類型試題所需項目，此表針對各種試題作分析，將試題可能的項目均考慮進來，並輔以試題參數用以適性測驗施測時預先填入試題參數。

表 2 各類型試題樣版所需項目

試題類型	所需項目
是非題	試題編號、題目名稱、題目、試題正確解答、正確回饋訊息、錯誤回饋訊息、試題參數。
單選題	試題編號、題目名稱、題目、試題正確解答、選項、對應選項回饋訊息、試題參數。
複選題	試題編號、題目名稱、題目、試題正確解答、選項、對應選項回饋訊息、試題參數。
排序題	試題編號、題目名稱、題目、試題正確解答、選項固定、選項、回饋訊息、試題參數。
填充題	試題編號、題目名稱、題目、試題正確解答、空格答案、對應空格回饋訊息、試題參數。
配合題	試題編號、題目名稱、題目、試題正確解答、配合主題、配合主題答案欄、試題參數。
問答題	試題編號、題目名稱、題目、正確回饋訊息、試題參數。

## 5. 結論

此章節將針對數位化題庫編輯模組的不足和缺陷來規劃未來題庫編輯模組可以擴充的方向，詳細內容如下：

1. 題庫系統試題的再利用：在題庫系統中，分析時並未考慮題庫與題庫之間試題的再利用，若要進行試題的再利用，需要先行將試題匯出，然後在匯入題庫之中，相當的不便，未來可以規劃試題的移動功能，也就是拖拉放(drag and drop)的功能，在新增或者是修改時能夠直接將別的題庫裡的試題複製一份到正在編輯的題庫。
2. 試題的重組：在目前題庫編輯模組中，並未考量到試題的重組，在未來的發展方向可以考慮增添此項機制，使用者只需要輸入測驗的目標，然後可以藉由系統資料庫的搜尋和重組機制，呈現新的試題給使用者，若要增加此項機制，必需先行建立測驗領域的知識本體，並將試題拆解解入各部分之下，若要重組試題，只需要到各個分類內抽取試題的題目或者是選項，就可以達到重組的功能。
3. 適性試卷的編製：在適性試卷編製時，需先制定需要的目標函數，然後針對此目標函數選出能夠最接近目標函數的試題，在試題的選擇時，需要大量的試題訊息[3]，現可以仰賴電腦的輔助，達到適性試卷的編製。
4. 試題反應模式的編輯：就目前的試題反應理論來說，存在著三種試題反應的模式，分別為單參數、雙參數、及三參數，但是除了這三種的試題反應模式之外，還是有一些與這三種不同的反應參數，在未來可以針對此項發展，讓題庫系統的使用者能進行反應模式的編輯，並且

能夠設定相關的參數值，增加題庫系統的靈活度。

## 參考文獻

- [1] 余民寧。1991。試題反應理論的介紹（一）：測驗理論的發展趨勢。研習資訊，8卷6期，第13-18頁。
- [2] 余民寧。1992。試題反應理論的介紹（三）：試題反應模式及其特性。研習資訊，9卷2期，第6-10頁。
- [3] 余民寧。1993。試題反應理論的介紹（十二）：電腦化適性測驗。研習資訊，10卷5期，第5-9頁。
- [4] 何榮桂。1991。電腦化題庫設計的基本架構。教育研究雙月刊，20輯，第21-26頁。
- [5] 何榮桂。2005。數位化題庫之概念架構。國家菁英季刊，1卷4期，第149-158頁。
- [6] 陳慶帆、許意苹、林敏慧。2005。MOODLE教學平台多媒體評量編輯系統之探討。台灣國際網路研討會，台中。
- [7] 鄒慧英。2002。測驗與評量—在教學上的應用。洪葉文化事業有限公司，第58頁。
- [8] Eduardo Guzmán and Ricardo Conejo, "A brief introduction to the new architecture of SIETTE," Adaptive Hypermedia 2004 Conference, Berlin, 2004.
- [9] Eduardo Guzmán and Ricardo Conejo, "Self-Assessment in a Feasible, Adaptive Web-Based Testing System," IEEE Transactions on education, Vol. 48, NO. 4, 2005.
- [10] H. Barbosa Len, H. Rego1, T. Moreira1, and F. Garca Pealvo1, "A model for online assessment in adaptive e-learning platform," The 3rd International Conference on Multimedia and Information & Communication Technologies in Education, Badajoz, 2005.
- [11] Smythe, Colin and Roberts, P., "An Overview of the IMS Question & Test Interoperability Specification," The fourth international computer assisted assessment conference, Loughborough University, 2000.