

磨課師學習分析框架—以 OpenEdu 為例

An Analysis Framework for MOOCs Learning-A Case Study in OpenEdu

陳仰萱 薛念林 陳錫民

逢甲大學資訊工程學系

Yang-Syuan Chen, Nien-Lin Hsueh, Hsi-Min Chen

Email: yangsyuan8012@gmail.com, nlhsueh@mail.fcu.edu.tw, seeme.goo@gmail.com

摘要

近幾年在數位學習帶來創新的磨課師，也就是大規模線上開放課程，在世界各地蓬勃發展。任何使用者只要透過網際網路即可存取到不須付費、高品質的課程內容。學習分析一直是數位學習當中十分重要的議題，既能夠瞭解學習者的行為，也能夠改善教材不足之處、因材施教等等。但是磨課師為學習分析帶來了新的挑戰，如資料的異質性與複雜性、目前磨課師資料分析的不足等等。

本文提出了一個磨課師資料分析框架，從 5W1H 六個不同的面向探討資料分析框架所需要的內容。並且為了驗證框架的有效性，實作了以 OpenEdu 為主的學習資料分析平台進行案例研究。任何磨課師的相關人員都可以透過分析平台獲取需要的資料分析內容。

關鍵字：線上教育、學習分析、學習分析框架、磨課師、學習行為

一、緒論

網際網路發展帶動了數位學習 (E-learning) 技術，即是利用資訊科技傳遞、擷取學習資訊的學習方式，其旨在於藉由所提供的數位學習內容及情境，創造有意義的學習經驗。近幾年興起的大規模線上開放課程 (Massive Open Online Course)，又稱磨課師，早在 2008 年提出，直到 2011 年才有大規模的發展。兩個關鍵特徵可以描述這個新的教育技術：開放式可存取性 (Open Accessibility) 與可擴展性 (Scalability)。磨課師的參與者不需要註冊一所學校或大學，也不用為參與課程而付費。磨課師提供的課程可能有預先制定的時間表並可能結合不同來源的影片、不同的線上資源、論壇等。在磨課師平台中，提供給學習者的資訊僅作為一個起點，學習者可以透過論壇與其他人溝通以找到解答和額外教材。磨課師的系統主要包含五個元素：教師、學習者、主題、教材和情境 (Context) [1]。目前全球最受歡迎的磨課師平台諸如 OpenEdX、Coursera、Khan Academy 以及 Udacity 等知名機構都提供近數百堂免費課程，任何人都能透過網際網路存取課程資源並與其他同儕互動。

在數位學習領域，改進或評估學習者的學習經歷持續地進行研究。在數位技術的帶領下，逐漸發展成一個新興領域稱為學習分析 (Learning Analytics)。事實上，學習分析並非屬於真正嶄新的研究領域，其涵蓋許多不同的相關領域並結合許

多既有的技術。學習分析在 2011 年第一次學習分析與知識國際研討會定義為「為了瞭解、優化學習與發生環境，進行測量、收集、分析與報告關於學習者的資料和他們的情境 (context)」[2]。具體來說，其旨在於瞭解遍布線上平台的學習過程，包含分析學習者、情境和最終與教學者之間的關係。學習分析通常為一個迭代的週期式處理過程，且包含三個部分 [2]：資料收集與前處理 (Pre-process)、分析與行動、後處理 (Post-process)。多數的學習分析數據是透過教育機構所建立的學習管理系統 (Learning Management System, 簡稱 LMS) 收集，是一種用來管理、文件化、追蹤、回報和傳遞電子教學材料的軟體。然而 LMS 經常僅作為支援教學的平台，成為教師與學生之間的通信工具 [3]。磨課師之所以受到重視，不只在於其「開放教育」的理念，也不只是不同於傳統教學的特色，而是在磨課師平台上大規模的線上學習行為 (Learning Behavior) 紀錄都完整地保留。

OpenEdX 為一開放原始碼專案，由麻省理工學院和哈佛大學創立。該系統核心 edX-Platform 包含學習管理系統 LMS 與課程創建工具 Studio。在國內，國立中央大學及逢甲大學團隊合作，基於繁體中文文化的 OpenEdX 共同推出「中華開放教育平台」，以下簡稱 OpenEdu。從 2014 年至今已有超過一萬名註冊使用者，建立近兩百門開放式課程，提供國內學習者良好的線上學習環境。

磨課師對學習分析帶來許多突破。舉例來說，這些知名的磨課師課程通常有數千名學生參與，但教師無法為每位學生奉獻許多時間。如此混亂的情況使得採用學習分析技術更加重要，但也可能揭露一些新的問題。除此之外，在磨課師平台上進行學習資料分析也有許多挑戰，如下所述。

● 異質性 (Heterogeneity)

由於磨課師的開放特性，其學習者來源相較於傳統教學方式更加廣闊，但有許多使用者最初加入是出自於好奇心，在註冊課程後便不再參與，僅少數學生本質上是真正地參與課程。學生群體的改變在磨課師產生的學習行為資料造成明顯不同的學習行為特徵且偏離常態分布。面對這樣不尋常的統計分布數據，將難以給予適當又客觀的建議。

● 複雜性 (Complexity)

磨課師平台幾乎都擁有龐大的使用者數量，因此也產生了可觀的網頁存取日誌 (Web Access Log)，代表學習者的學習行為。在學習分析中，磨課師的資料包含了日誌 (Log)、背景訊息以及

相關的課程內容。資料的多維度特性，系統嚴格地需要處理大量且多來源、多格式、非結構化的資料，在處理與分析上都帶來許多的困難。

目前有許多知名磨課師機構也致力於學習分析。OpenEdX Insights 使課程教師或管理角色聚焦於單一課程，提供如學生參與活動次數、學生背景分布與學習效率狀況，以改善教師之課程設計，如下圖 1。



圖 1 OpenEdu Insights

但圍繞在磨課師周圍的利益關係人所需要的學習分析資訊大不相同，舉例來說，若是平台的管理者，則希望掌握所有課程的使用狀況與學習效能。由於日誌詳細記錄學習的生態，因此每位角色都能透過分析取得有用的資訊。在數位學習的發展中，評估學習效能的學習指標（Indicator）已有相關研究[4]，但目前 Insights 所提供的資訊十分不足，並且由於資料的特性，也難以直接應用現有的學習指標計算方法。

本研究基於 OpenEdu 學習資料，採用 5W1H 分析法，又稱六何分析法，從不同角度探討磨課師學習分析框架所須具備的特性，分別是使用對象（Who）、分析目的（Why）、分析指標（What）、透過何種技術進行分析（How）、提供結果的時機（When）以及最後於何處交付分析結果（Where），如下圖 2 所示。適用於採用 OpenEdX 平台作為磨課師課程提供者，欲進行資料分析時皆可應用此框架，以幫助調查圍繞在磨課師周圍的學習成果與評估，並揭露任何可以用於學習優化的有用資訊。

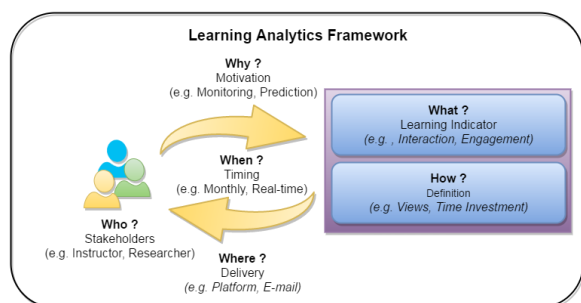


圖 2 Learning Analytics Framework

● 角色（Who）

磨課師周圍有許多不同的利益關係人，例如提

供平台的營運者、實際進行教學的教學者、參與學習的學習者等等，這些角色都對於學習分析有不同需求。分析框架必須辨認這些不同的角色以提供不同的分析服務。

● 目的（Why）

根據磨課師周圍角色不同的觀點產生的學習分析需求是不同的，例如監測、分析、預測、干預、評估、適性化、個人化、推薦和反饋等等。因此分析框架需要能夠完整地解讀原始資料，提供符合各個利益關係人需求的分析結果。

● 分析指標（What）

分析框架須要量身定制一套指標與衡量標準以滿足多樣化的需求。例如以適當的指標衡量學習者的效能，或是學習行為指標警示教師及早干預學習者。其中的挑戰在於制定在某個情境中的分析指標與度量標準。並且需求會不斷改變，資料分析框架必須擁有彈性的可擴展性。

● 分析方法（How）

適當的學習指標還需適當的計算方法，多數學習分析皆採用學習行為歷程計算取得。其分析方法多樣化，例如機器學習、人工智慧、資訊檢索、統計等等。並且為了要讓獲取分析結果的使用者易於理解，通常會採取資訊視覺化（Information Visualization）方式呈現。除此之外，教師或教育研究者對於學習資料可能有不同的分析方法，因此分析框架需要提供可供再利用、可理解的學習資料。

● 時機（When）

資料提供時機主要取決於角色的需求與系統處理時間。因應不同需求，例如教師可能會想在學期開始前先瞭解學生群體的背景訊息，而在課程中期瞭解學生的學習情況，以便中途干預或是修正教學內容。但是若以教育研究者而言，對於學習行為資料需求的急迫性便低於課程教師。資料分析框架必須在適當的時機提供符合角色需求的分析結果。

● 交付（Where）

最後，總結上述幾個面向，必須要有一個系統或方法能夠在適當的時機交付適當的結果給適當的角色。處理與應用學習分析是需要低延遲的，鑒於不同角色對於可接受的計算時間都不同，例如研究者可能願意接受幾分鐘甚至數小時而教師可能想要只有幾秒的即時狀況。因此，學習分析框架將需要依照這些特性所建構。

本研究的目的是在於探索上述六個面向並提出一個學習分析框架，能夠測量、收集、分析並回報學生的行為，以改善 OpenEdu 資料分析的現況。最後，為了驗證其框架的有效性，實作了一個以 OpenEdu 為主的學習資料分析系統，提供給不同課程角色所需的相關學習分析資訊。

二、相關研究

自從被稱為磨課師元年的 2012 年以來，國內仍然少有關於磨課師學習分析之相關文獻，而國外

已有許多相關研究產出。在此方面的研究大致分為兩個類別，其一是提出或開發一個軟體架構以收集、分析磨課師學習資料；其二是針對現有磨課師課程內容進行學習分析。

2.1 軟體架構方面

在數位學習領域中，各式各樣的軟體平台蓬勃發展，因此資料的複雜性與異質性一直是進行學習分析的挑戰。Kagemann 等學者使用語意 (Semantic) 技術建立磨課師的語意資料模型，應用鏈結資料 (Linked Data) 的方式建立一個統一資料收集的網站 MOOCLink[5]，藉此整合不同平台的資料，提供開放課程的比較與探索。

隨著各式各樣的磨課師平台出現，也有一些研究專注於改善提供課程的平台分析功能。Ruipérez-Valiente 等學者 [6] 針對可汗學院 (Khan Academy) 開發擴充的分析平台。包含整個課程和個別學生的全新視覺化，基於之前的研究 [7] 包含 21 個指標，以瞭解平台的使用狀況與學習效能。Ruiz 等學者 [8] 在 OpenEdX 平台上建立一個擴充資料分析套件。教師可以利用視覺化查看課程整體進度、監控每位學生作業進度與花費時間；學生則是能獲得學習過程進度的自我反饋。但目前這些平台提供的學習行為資訊深度及廣度都略嫌不足，並且僅針對教師及學生提供分析結果，並未考量其他磨課師角色也有資料分析的需求。

磨課師學習分析所需的資料除了複雜且難以處理以外，也要考慮如何處理學習資料的生命週期。因此，有許多研究提出了相應的軟體架構進行學習資料分析。較早的 Santos 等學者 [9] 開發一個視覺化儀錶板顯示學生的參與程度並可與同儕比較，並且在研究中進行了四次的迭代改進其易用性 (Usability) 與可用性 (Availability)。最後評估表明儀錶板對於學生與教師是有用的。Shi 等學者 [10] 則開發 VisMOOC 透過視覺化的方式觀察學習者的影片點擊流 (Clickstream) 行為，以多種視角幫助教師分析學習資料。Tabaa 等學者 [1] 建立一個以 Hadoop 為基礎的平台收集、分析平台所記錄的大量資料。

完整記錄了學習行為的日誌資料是學習資料分析中不可或缺的資料來源，Zheng 等學者 [11] 將數位學習的系統依據可能產生的日誌來源分為四層：基礎層 (Infrastructure Layer)、內容層 (Content Layer)、應用層 (Application Layer) 與終端層 (Terminal Layer)，並提出一個架構以分析這些教育日誌資料。Zheng 等學者所提出的架構較為通用，適合進行多平台的收集分析，或許可以作為日後應用的基礎。

OpenEdX Insights 同樣是為了學習分析所開發的。下圖 3 為 Insights 系統架構，左側 LMS 包含學習分析資料如記錄學習者互動的 Event Tracking 與 State，程式會週期性地壓縮並複製這些檔案至檔案系統供 Pipeline 存取。OpenEdX

Analytics Pipeline 讀取 LMS 產生的資料經過處理並分析後，獲得的結果儲存於 Result Store。其中 Scheduler 排程 Pipeline 執行資料計算任務以更新 Result Store 的資料。最後，Application 的 Insights 透過 edX Analytics Data API 呈現資料。使用者透過支援的網路瀏覽器存取資料，Insights 直接與 LMS 溝通驗證並授權使用者讀取課程結構相關訊息。為了因應大量學習者所產生難以估計的學習行為資料，主要進行學習分析的 Pipeline 基於 Hadoop Hive 和 MapReduce 執行資料分析任務。

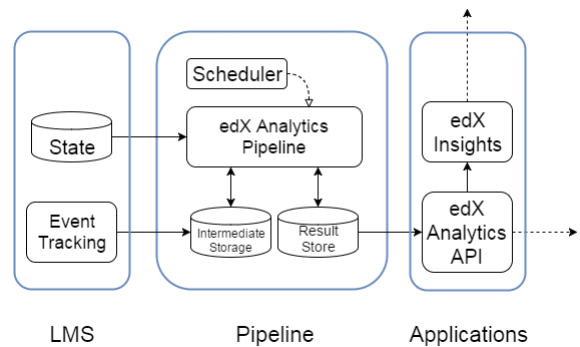


圖 3 Insights 系統架構

Insights 所提供的學習分析內容並不十分完善，並且僅針對教師及合作院校提供。而 Insights 為一個開放原始碼專案，更新頻繁且相關文件不足，因此在 Insights 的分析功能擴充上仍有許多困難與挑戰。

2.2 資料分析方面

磨課師發展早期的學習分析研究著重在描述性統計。DeBoer 等學者 [12] 研究結果顯示學生的學習效率或成績與學生的背景資訊有顯著的變化。Seaton 等學者 [13] 則是研究 MITx (現今的 OpenEdX) 的第一門磨課師課程學習行為，除了解析日誌、分析各個獨立課程資源的描述性統計與總時間花費以外，也發現學習者在解決作業與考試的行為上有些許差異。

磨課師平台的使用者來源十分廣泛，因此在課程中可能有許多對於該課程瞭解程度不一的學習者。Vasić 等學者將學習者依據布魯姆分類法 (Bloom's taxonomy) 將學生的知識水平 (Outcome Level) 分為五大類別 [3]，再依據學習者的行為透過分類模型預測學生的知識水平。儘管實驗顯示其預測模型相當準確，但僅基於有限的訓練資料，且限於特定的技術相關課程進行實驗。

由於磨課師的學習成本低，造成學習者很容易放棄繼續學習。因此磨課師課程中大量的註冊學習者與極少的完成課程人數形成強烈對比，極低的完成率成為了許多研究探討的議題。Jordan 透過社群的方式收集資料，採用線性回歸分析 91 門課學生的選課及完成率 [14] 之間的關聯性。發現選課人數與課程長度有關係，完成率則無。完成率通常與

開課時間、大學排名、選課人數有關，而且與課程長度反向相關。Cisel 則針對法國一門磨課師專案管理課程進行分析，探討學習者的成就達成與哪些因素有關[15]，研究指出成就是否達成與他們的工作狀態、區域、及時間限制有很大的關係。另外，研究也顯示出討論活躍且有進行同儕互評（Peer Review）能夠完成成就者較多。

除了探討課程完成率的研究外，Halawa 等學者試圖找出何種行為的學生更容易放棄學習，亦即輟學（Dropout）[16]。作者透過建立分類預測模型發現造成放棄學習的因素有自我效能感知（Self-perceived Self-efficacy）低、自我調適能力（Self-regulation）低、及興趣低。研究表明預測模型仍在學習的學生中，也能找出 40 % 至 50 % 可能退出的學生，而超過兩個星期缺席的學生，則能找出將會輟學的 90 % 學生，使得教師或平台能夠及早進行干預。

如上述，多數研究著重於優化學習者的學習行為，但教材同樣為磨課師重要元素之一，而影片則是線上教育廣泛使用的素材。Guo 等學者[17]分析影片製作方式如何影響學習者的參與，針對 OpenEdX 平台四門課程近七百萬觀看次數（video watching session）進行分析，探討學習者於不同製作方式的影片觀看情況。研究發現影片長度越短、類似 Khan 方式的平板書寫解說方式與教師說話速度較快且具備高度熱情的影片，學習者的參與情況越好。在磨課師課程中，學習者通常倚賴討論區與其它學習者互動。Coetzee 等學者[18]針對 OpenEdX 中 1,101 位學習者進行田野實驗（Field Experiment），探討信譽系統（Reputation System）對於論壇行為的影響。研究發現信譽系統能產生更快速的回覆時間與大量的回應。然而，在學習成果與社群主觀意識上則無顯著影響。研究結果意味著論壇對於磨課師是必不可少的，改善學習成果與形成社群則仍需要其它技術。由此可見，學習分析除了提升學習者的學習效能外，也能從中發掘對教師設計課程教材有益的資訊。

三、研究方法

磨課師產生的學習數據資料生命週期為獲取資料（Acquire Data）、組織（Organize）、分析（Analyze）與優化學習（Optimize Learning）[1]。「獲取資料」在於從資料來源週期性地擷取資料。「組織」則代表異質多來源的資料需要進行適當的前處理與轉換，準備進行深入的分析。在「分析」階段，每位磨課師角色對於學習資料分析有不同的需求，為了滿足不同的需求，建構適當的分析框架是很重要的。最後，呈現給所有角色的分析階段的「結果」，使不同層面的人員皆能獲取想要的資訊，以優化學習環境與過程。

本研究提出的方法也將覆蓋完整的磨課師資料生命週期，而處理學習分析的流程則如下所述，每個處理程序都關聯於圖 2 中的每個面向。

3.1 測量

學習平台必須運用各種方法，以統一的標準測量學習行為。本研究以 OpenEdu 作為分析對象，在此介紹平台所記錄的相關資訊，包含結構化（Structured Data）的使用者資訊與半結構化（Semi-structured Data）的課程資訊與學習行為日誌。

● 學生資訊

使用者註冊 OpenEdu 時便要求填寫背景相關訊息，如性別、教育程度等，並儲存於關聯式資料庫中，如下表 1。

表 1 學習者背景資料（範例）

Id	User Id	Name	Gender
1000	999	Open Edu	m
Year of Birth	Level of Education	Country	City
2015	professor	TW	Taichung

● 課程資訊

課程結構在 OpenEdu 當中以一種輕量的文本型格式 JSON（JavaScript Object Notation）儲存於非關聯式資料庫中。包含教師定義的課程結構內容、課程的基本資訊還有課程討論區中的言論內容。

● 學習行為

最重要代表學習行為的便是系統所記錄的詳細日誌資料，同樣以 JSON 格式儲存，如下表 2 是一個影片播放日誌的部分範例。主要以事件驅動（Event-driven）方式記錄學習者對於課程元件中的每一個點擊行為，例如點擊播放影片、提交問題答案等等，也包含使用者本身的訊息與互動元件相關的訊息。

表 2 一位學生點擊播放影片 Log（部分）

{
"username": "OpenEdu",
"ip": "123.123.123.123",
"event_type": "play_video",
"event": {
"currentTime": 243,
"code": "iOOYGgLADj8",
"time": "2014-12-23T14:26:53.723188+00:00"
}
}

3.2 收集

不同角色（Who）需要不同的學習分析（Why）。學習分析系統必須收集上述所需的資料並進行適當的前處理，萃取出易於分析處理的資料與保存格式（How）。在 OpenEdu 日誌資料前處理當中發現了以下問題：

● 資料的有效性（Validity）

日誌訊息包含互動物件的資訊，例如物件 ID、存取物件的網址等等。可能是保存的資料超過了固定的資料長度，造成日誌記錄不完整，因此並非一個正確的 JSON 格式，而這樣的資料必須在

前處理先行過濾。

- 資料的真實性 (Veracity)

每個日誌會記錄日誌的產生時間，一個低延遲的系統中，同時也代表學習者與課程元件的互動時間。初步實驗觀察 OpenEdu 日誌的過程中發現不合理的連續影片播放日誌。因此這樣的行為也必須在分析的過程中過濾。

- 資料的組織方法 (Organization)

OpenEdu 所進行的日誌機制以時間序列為基準，包含整個平台的使用者互動行為。但在進行資料分析時，需要關注的是單一使用者行為以瞭解特定的操作模式。因此，為了易於分析必須先進行資料切割，將時間序列的日誌轉換為單一使用者的日誌組織方式。

3.3 分析

如下表 3 因應不同角色 (Who) 產生不同需求 (Why)，因此而需要採用不同的學習分析模型 (How) 計算分析指標 (What)，如下表 4 所示。

表 3 不同利益關係人所需求的學習分析項目

角色	需求
平台 (Funder)	學習者分布、課程使用率、學習成效、教師營運狀態
教師 (Instructor)	學生的學習效能評估、教材使用率
學習者 (Learner)	學習狀況指引、個人化推薦
研究者 (Researcher)	學生的學習風格、學習模式
公眾 (Publics)	公開可存取的磨課師開放資料

表 4 各種不同的分析項目

	影片	作答	討論	課程
常見分析	使用者背景資料、存取頻率、資源利用率、時間投資			
特定分析	加、減速觀看、重播、倒轉行為	作答正確率、作答分布	內涵分析、意見領袖 (Opinion Leader)	不同類別課程指標比較
輟學分析	與學習行為之間的關聯			
取得證書				

本研究將針對以下幾個面相的學習分析：

- 學習者行為

磨課師課程中提供的多樣化資源，如影片、作答與討論區等等，透過存取行為發掘教材可改善之處。除了學習者面向的行為分析，對於平台管理者來說，教師的經營過程也同樣重要。一門相同內容的課程是否有教師或助教與學生互動，不僅在註冊人數和取得證書的人數上都有顯著的差距。因此，呈現教師與學生的互動過程能夠提供平台管理者

更多資訊以評估授課教師。由於時間是學習者與教師僅能投入的資源，因此一個使用者投入的時間越多，也代表平台的使用效率越高。

磨課師的開放可存取性使得學習者來自四面八方。本研究除了呈現學習者的背景資訊分布外，也同樣探討不同背景變項之學習者的學習行為是否不同。此外，也能夠讓平台或是教師了解課程受眾類型以應用於宣傳或是調整授課內容。

- 輟學分析

由於磨課師也屬於遠距教學的一種，學習者能夠輕易的放棄學習。以 OpenEdu 來說，學習者除了註冊課程外還能夠自由取消註冊課程，但是一個放棄學習的學習者更可能的是再也不進入該平台。因此「輟學」除了包含取消註冊的學習者以外，也應包含一段時間未進入課程的學習者。這樣的訊息能夠幫助教師與平台即時干預，或是透過進一步分析輟學者的學習行為以瞭解放棄原因。

- 學習成果分析

OpenEdu 定義的「學習成果」為是否取得修課證書，由於許多學習者能力的差異，不見得所有學習者都會取得修課證書。除此之外，前述所提到的學習者異質性的挑戰，透過過濾出輟學的學習者，其餘的使用者才是真正想要或完成課程的學習者。由於他們才是消耗平台資源的最大使用者，因此，必須要將觀點放在這些活躍學習者的學習行為，所得到的學習效能評估將能夠更加準確。

本研究除了統計學習者背景資料以外，也透過演算法計算使用者點擊課程元件之間的時間，以作為該學習者使用此課程元件的投入時間。每個分析項目都需要採用不同的分析方法或指標。

3.4 回報

除了前述提到的不同角色 (Who) 以外，還要考量呈現時機 (When)、如何呈現 (How) 與在何處 (Where)。如下表 5 舉例各個角色對於需求分析可能希望瞭解的時機，而框架也應於合理的等待時間內提供分析結果。

表 5 各角色希望取得分析之時機

角色／資料	學習者分布	課程使用率	學習成效	公開資料
平台 (Funder)	營運期間、每月	營運期間、每月	營運期間、課程結束時	／
教師 (Instructor)	課程開始時	開課期間、每週	開課期間、每週	課程結束時
學習者 (Learner)	／	／	開課期間、每天	／
研究者 (Researcher)	課程結束後			
公眾 (Publics)	營運期間、年度			

研究提到以對使用者友善的圖形取代表格形式的

統計資訊將有助於解釋教育數據的分析[2]。因此一個良好的分析框架應以使用者容易理解的視覺化資訊呈現。呈現結果的系統架構則於下一章節進行說明。

四、系統設計

本研究將實作上述研究方法並開發 OpenEdu 資料分析系統，下圖 4 為系統架構。主要分為四個層面並使用 R 語言進行實作。R 語言是一種開放原始碼的直譯程式語言與操作環境，有強大的統計分析、資料探勘與繪圖能力。

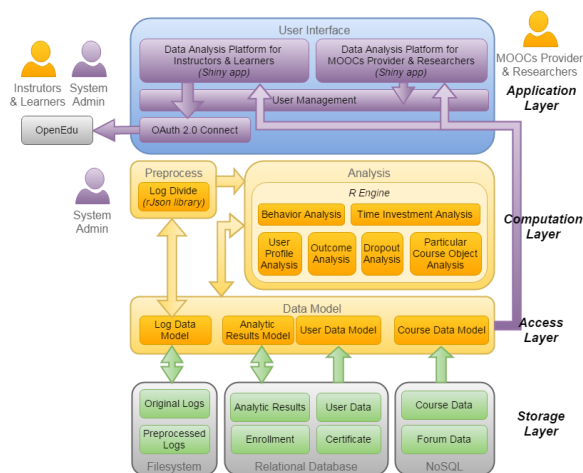


圖 4 系統架構

以下將介紹系統架構四個層面所包含的模組，4.1 節描述與使用者互動的應用層，4.2 節為資料分析的計算層，4.3 節為以利資料存取的存取層，最後 4.4 節則是說明不同的資料儲存架構。

4.1 應用層

應用層提供不同利益關係人 (Who) 各自不同的互動式介面。下圖 5 為開發中的平台畫面，除了呈現資料分析結果也提供學習資料下載再利用 (How)。以下將詳細說明相關的模組。

- 資料分析平台：透過存取層的資料 Model 模組取得不同的資料分析結果呈現。
- 使用者管理：與 OpenEdu 驗證會員資料，採用 OAuth 2.0 連線模組進行溝通以因應不同的使用者呈現不同的分析結果。
- OAuth 2.0 連線：OAuth 2.0 是一個開放標準允許第三方應用程式在安全的情況下存取資料。

4.2 計算層

計算層提供不同角色 (Who) 所需的各種不同分析目標 (Why)，為了因應提供資料的時機 (When)，分析模組以許多不同週期迭代式產生分析結果 (How、What)。首先，學習分析資料首先需要前處理才能產生易於分析的資料格式。接著分析模組如 3.3 節所述有許多不同的分析內容，透



圖 5 OpenEdu LA 開發中畫面

過資料 Model 模組進行資料存取，以許多不同的方法 (How) 分析後將計算結果儲存至儲存層的關聯式資料庫。

- 前處理：為了不同的分析目的，需要將資料轉換為容易分析的資料作為輸入。本研究為了以使用者為分析對象，將學習行為日誌轉換為以使用者為主的組織方式。
- 分析：透過文獻回顧探索應用於磨課師平台的學習分析指標，進行符合各角色需求的學習分析。每一個學習分析項目都需要一個特殊的演算法進行計算。

4.2 存取層

磨課師產生的學習資料十分多樣，學習分析計算 (How) 需要頻繁的資料輸入及輸出，因此需要定義資料存取介面以便安全、有效地存取資料 (What)。為了使系統更有彈性地取用所需的資料，存取層為與其它模組互動的重要模組。

- 日誌資料 Model：前處理模組透過此模組取得原始日誌且輸出處理後的日誌至檔案系統。
- 分析結果 Model：提供資料分析平台呈現分析結果，也有一些分析項目會從分析後的結果再次進行不同的分析計算。
- 使用者與課程資料 Model：學習分析與學習者及課程內容息息相關，此模組提供平台及分析模組所需的學習者及課程結構資料。

4.3 儲存層

儲存層保存原始資料以及存放運算後的結果 (What)。本研究中，其他模組透過資料 Model 模組與不同的儲存結構互動。

- 關聯式資料庫：儲存學習者背景、註冊課程

之相關訊息以及分析結果。

- 非關聯式資料庫：儲存課程與討論區資料。
- 檔案系統：儲存每日遞增的使用者行為日誌，包含經過前處理的日誌資料。

目前的 OpenEdX 平台除了 Insights 之外，本身的 LMS 僅提供一個基礎頁面給學習者查看在課程中的整體進度，另外身為課程教師的使用者則有另一面板能夠查閱課程註冊人數。因此，在使用者的互動平台上，學習分析功能仍是十分短缺。在 2.1 節描述了 Insights 的介面示意圖及架構，但目前該專案仍然處於開發中版本（Version 0，Alpha），因此提供的資料分析支援及相關文件仍然不足且有頻繁地變更。

本研究所提出的架構包含應用層、計算層與儲存層。包含磨課師學習資料收集、轉換、儲存、計算與提供服務。並且每個層面都必須符合考量的 5W1H 面向，不僅是服務學習者與教師，而是呈現最適學習分析結果給磨課師的周圍利益關係人。

五、案例研究

本研究針對 OpenEdu 探討使用者之學習行為分佈。

表 6 簡述每個模組所包含之分析指標

分析模組		學習分析指標
學習者背景資料		平均年齡、性別：註冊平台的學習者背景訊息分佈。
投入時間		課程活動時間分佈：使用者觀看影片、作答、討論的時間區間。
輟學分析		輟學率：輟學者占總修課人數之比例。 輟學與行為分析：輟學者觀看影片、作答、討論的行為分佈。
課程資源存取		修課人數：一門課程自開課至結束時所有註冊的使用者數量。 完課率：取得修課證書占修課人數之比例。 完課與行為分析：完課者觀看影片、作答、討論的行為分佈。
	影片	觀看次數、人數：每部影片之瀏覽超過三秒之人次、人數。 觀看率：每位學習者觀看影片數量占總影片數量比例。
	評量	作答次數、人數：評量提交答案之人次、人數。 作答率：每位學習者作答題目數量占總題目數量比例。
	討論	發表次數：每位使用者於討論區發表文章次數。 討論率：討論人數占課程修課人數之比例。
特定課程資源分析	影片	觀看完整率：每部影片被完整觀看的比例。 快轉率：每部影片採用快轉或倒轉觀看影片的比率。
	評量	正確率：每位學習者作答題目之正確率。
	討論	參與程度：評量討論字數為討論參與程度。

資料來源為 OpenEdu 學習者背景訊息、課程內容、討論內容及學習行為日誌相關資料，從 2014 年 9 月 11 日至 2016 年 3 月 31 日進行收集，橫跨 103 及 104 學年度的磨課師課程。其資料量分別為學習者相關資料 628 MB、課程及討論內容 1.38 GB 及學習行為日誌 8.18 GB。並且關注其中 30 門課程進行重點案例分析。

如上表 6 簡述分析模組所包含之學習分析指標，這些指標將因應不同的使用角色而有不同的呈現。舉例來說，平台管理者之主要需求為瞭解平台使用狀況與各開課教師之營運狀況。因此，分析項目將著重於課程與教師之間的比較，以及呈現所有課程的概況。教學者之主要需求為瞭解學習者學習效能與課程教材使用狀況。因此，分析內容將會高度關聯於課程進行日期，以利教師分析每週課程之進度與學習情況。。

5.1 有效使用平台

包含表 6 的學習者背景資料、投入時間與課程資源存取三個項目都屬於平台的整體使用行為，從全面性的角度瞭解平台課程中的使用情況。

5.2 高效使用平台

磨課師課程的最大特點便是透過影片、作答與討論進行學習，本節描述這些特定的課程資源能夠進行的個別特殊分析項目，以瞭解使用者在平台上是否有效率地進行學習或是授課。

5.3 輟學分析

輟學定義為在課程結束前，學習者於課程結束前已連續三週未曾再進入課程或是主動選擇取消註冊課程之學生。輟學分析需要將輟學者的學習行為如上述觀看影片、作答與討論等等與其他學習者的行為進行區分，將更有助於瞭解輟學者的學習行為分佈，以利教師或平台即時進行干預。

5.4 學習成果分析

學習者進行評量所得到的成績取決於教師的課程設計而有所不同，並且由於 OpenEdX 證書系統對於非英語系的不支援，而無法取得學習者的課程總成績。因此本研究中 OpenEdu 學習成果僅以取得課程證書的學習者為基準。學習成果分析也將取得證書的完課學習者行為與其它學習者行為進行區隔，以瞭解完課者的學習行為分佈。

六、結論

磨課師的發展為學習分析帶來許多挑戰，本論文透過 5W1H 分析法提出一個適用於磨課師平台的資料分析框架，從磨課師平台周圍利益關係人（Who）、不同的分析需求（Why）、如何分析（How）相關指標（What）以及最後在何處（Where）交付分析結果，多個不同的角度探討磨課師資料分析框架所需要考慮的面向。

本文提出的架構包含四個層面，因應不同角色

的需求特性設計其分析與結果呈現。為了驗證框架的有效性，本研究使用 OpenEdu 學習資料進行系統開發及案例研究。這些資訊使得使用者能夠透過許多非原本進行學習的系統所提供的訊息輔助決策制定。

雖然學習分析在數位學習領域中十分重要，但所呈現的內容仍高度相關於多樣化的課程設計內容。仍然必須提供適當的學習分析資料使教師或教育學者進行各自需求的特殊分析內容。未來則會應用更多學習指標分析學習者的學習模式並評估其學習效能，而本系統對於學習效能上的影響則需進一步驗證。

七、致謝

本研究承蒙國家科學委員會研究計畫編號：MOST 104-2627-M-035-007 之計劃經費補助。

參考文獻

- [1] Y. Tabaa and A. Medouri, "LASyM: A Learning Analytics System for MOOCs," *Advanced Computer Science and Applications*, pp. 113-119, 2013.
- [2] M. A. Chatti, A. L. Dyckhoff, U. Schroeder and H. Thijs, "A Reference Model for Learning Analytics," *International Journal of Technology Enhanced Learning*, pp. 318-331, 2012.
- [3] D. Vasić, A. Pinjuh, M. Kundid and L. Šerić, "Predicting student's learning outcome from Learning Management system logs," in *Proceedings of the 2015 23rd IEEE International Conference on Software, Telecommunications and Computer Networks (SoftCOM)*, 2015.
- [4] A. L. Dyckhoff, V. Lukarov, A. Muslim, M. A. Chatti and U. Schroeder, "Supporting Action Research with Learning Analytics," in *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, 2013.
- [5] S. Kagemann and S. Bansal, "MOOCLink: Building and Utilizing Linked Data from Massive Open Online Courses," in *Proceedings of the 2015 IEEE International Conference on Semantic Computing (ICSC)*, 2015.
- [6] J. A. Ruipérez-Valiente, P. J. Muñoz-Merino, D. Leony and D. C. Kloos, "ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform," *Computers in Human Behavior*, pp. 139-148, 2015.
- [7] P. J. Muñoz-Merino, J. A. R. Valiente and C. D. Kloos, "Inferring Higher Level Learning Information from Low Level Data for the Khan Academy Platform," in *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, Leuven, Belgium, 2013.
- [8] J. S. Ruiz, H. J. P. Díaz, J. A. Ruipérez-Valiente, P. J. Muñoz-Merino and C. D. Kloos, "Towards the Development of a Learning Analytics Extension in Open edX," in *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*, 2014.
- [9] J. L. Santos, S. Govaerts, K. Verbert and E. Duval, "Goal-oriented Visualizations of Activity Tracking: A Case Study with Engineering Students," in *{Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge}*, 2012.
- [10] C. Shi, S. Fu, Q. Chen and Q. Huamin, "VisMOOC: Visualizing Video Clickstream Data from Massive Open Online Courses," in *Proceedings of the 2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*, 2014.
- [11] Q. Zheng, H. He, T. Ma, N. Xue, B. Li and B. Dong, "Big Log Analysis for E-Learning Ecosystem," in *Proceedings of the 2014 IEEE 11th International Conference on e-Business Engineering*, 2014.
- [12] J. DoBoer, G. Stump, D. Seaton and L. Breslow, "Diversity in MOOC Students' Backgrounds and Behaviors in Relationship to Performance in 6.002x," *Teaching and Learning Laboratory*, p. 16, 2013.
- [13] D. T. Seaton, Y. Bergner, I. Chuang, P. Mitros and D. E. Pritchard, "Who Does What in a Massive Open Online Course?," *Communications of the ACM*, vol. 57, no. 4, pp. 58-65, 2014.
- [14] K. Jordan, "Initial Trends in Enrolment and Completion of Massive Open Online Courses," *The International Review Of Research In Open And Distributed Learning*, pp. 133-160, 2014.
- [15] M. Cisel, "Analyzing completion rates in the First French xMOOC," in *Proceedings of the European MOOC Stakeholder Summit 2014*, 2014.
- [16] S. Halawa, D. Greene and J. Mitchell, "Dropout Prediction in MOOCs using Learner Activity Features," in *Experiences and best practices in and around MOOCs*, 2014.
- [17] P. J. Guo, J. Kim and R. Rubin, "How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos," in *Proceedings of the First ACM Conference on Learning @ Scale Conference*, 2014.
- [18] D. Coetzee, A. Fox, M. A. Hearst and B. Hartmann, "Should Your MOOC Forum Use a Reputation System?," in *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 2014.