

# 短文回應的主題自動歸類在 行動教育活動上之應用初探

A Preliminary Study of Short Text Responses Clustering  
for Mobile Educational Activities

李龍豪

**Lung-Hao Lee**

國立臺灣師範大學圖書資訊學研究所博士後研究員  
Postdoctoral Fellow  
Graduate Institute of Library and Information Studies  
National Taiwan Normal University

簡佑達

**Yu-Ta Chien**

國立臺灣師範大學科學教育中心博士後研究員  
Postdoctoral Fellow  
Science Education Center  
National Taiwan Normal University

張俊彥

**Chun-Yen Chang**

國立臺灣師範大學科學教育研究所講座教授  
Chair Professor  
Graduate Institute of Science Education  
National Taiwan Normal University

---

通訊作者：曾元顯 [samtseng@ntnu.edu.tw](mailto:samtseng@ntnu.edu.tw)

投稿日期：2016年11月6日；接受日期：2016年12月15日

Email：李龍豪 [lhlee@ntnu.edu.tw](mailto:lhlee@ntnu.edu.tw)；簡佑達 [yutachien@ntnu.edu.tw](mailto:yutachien@ntnu.edu.tw)；張俊彥 [changcy@ntnu.edu.tw](mailto:changcy@ntnu.edu.tw)；  
李宗諺 [yan@ntnu.edu.tw](mailto:yan@ntnu.edu.tw)；曾元顯 [samtseng@ntnu.edu.tw](mailto:samtseng@ntnu.edu.tw)

李宗諺

**Tsung-Yen Li**

國立臺灣師範大學科學教育中心專任研究助理

Research Assistant

Science Education Center

National Taiwan Normal University

曾元顯

**Yuen-Hsien Tseng**

國立臺灣師範大學圖書資訊學研究所教授

Professor

Graduate Institute of Library and Information Studies

National Taiwan Normal University

## 【摘要 Abstract】

文字歸類是用來偵測文件語料涵蓋的主題，以便於資訊瀏覽、分析、組織的資訊檢索技術。而近年教室中即時反饋系統（Instant Response System, IRS）的運用越來越普遍，以強化學生專注參與的情形，進而增進其學習成效。然而目前的 IRS 較缺乏學生自由回應的簡短文字之歸類處理功能，以致無法進一步發揮其功效。本文的目的，在探討 IRS 現況以及短文回應的特性與問題、回顧短文主題歸類自動處理的技術與應用情形、提出一套可行的搭配系統供 IRS 整合運用，並展示我們目前實作的結果範例，以做為後續運用、研究的參考。本文比較了三種自動歸類方法，結果顯示影響歸類成效的因素眾多，理論上較好的歸類方法，不見得會有較佳的結果，還需其他因素或處理流程的配合。

Text clustering is a powerful information retrieval technique to detect topics from document corpora, so as to provide information browsing, analysis, and organization. On the other hand, the Instant Response System (IRS) has been widely used in recent years to enhance student engagement in class and thus improve their learning

effectiveness. However, the lack of functions to process short text responses from the IRS prevents the further application of IRS in classes. Therefore, this article aims to discuss the current status of IRS application, illustrate the characteristics of short text responses from the IRS, review current short text processing techniques, propose a proper short text clustering module for the IRS, and demonstrate our implemented techniques through real-world examples, so as to provide experiences and insights for further study. In particular, we have compared three clustering methods and the result shows that theoretically better methods need not lead to better results, as there are various factors that may affect the final performance.

### 【關鍵字 Keywords】

文字歸類；短文；即時回應系統；資訊組織；行動教育  
Text Clustering; Short Texts; Instant Response Systems; Information  
Organization; Mobile Education

## 壹、前言

文字歸類 (Text Clustering) 是傳統用來偵測文件語料涵蓋的主題，以便於資訊瀏覽、分析、組織的資訊檢索 (Information Retrieval, IR) 技術或是自然語言處理 (Natural Language Processing, NLP) 技術。隨著數位文字越來越多樣化，以及相關的應用越來越豐富，文字歸類技術的發展與應用，也需隨之改變與適應。

近年來，資訊技術領域的 IR 與 NLP 進展之快，已經對學術界與產業界產生深遠的影響。而教育領域的環境、工具與教學互動，也進步到對全球社會產生重大衝擊，例如：大量開放線上課程 MOOCs (Massive Open Online Courses)<sup>1</sup> 與課堂反饋系統 (Classroom Response Systems, CRS) (Siau, Hong, & Nah, 2006) 等。此兩個原先不同的領域已有越來越多融合之處，以支援過去無法進行的線上或課堂的教育活動，例如即時瞭解學生反映 (Beatty & Gerace, 2009) 以及行動化的語言學習 (Cardoso, 2010)。

---

1 [http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html?\\_r=0](http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html?_r=0), accessed on 2016/10/13.

這類的研究，越來越受重視（Hearst, 2015）。在學術界，美國地區自 2003 年起舉辦 NLP 建構教育應用創新工作坊（Workshops on Innovative Use of NLP for Building Educational Applications, NLP-BEA）<sup>2</sup>，而亞洲地區自 2014 年起舉辦 NLP 在教育的應用工作坊（Workshops on Natural Language Processing Techniques for Educational Applications, NLP-TEA）<sup>3</sup>，前者著重於英文或西方語文，而後者著重於中文或亞洲語文，兩者都是將文字處理技術應用於教育的學術活動。

在產業界，線上寫作評估系統如 ETS Criterion<sup>4</sup> 以及論文抄襲偵測（或文章重複性辨識）系統如 Turnitin<sup>5</sup> 都建立起相當龐大的市場。然而，這些成功的服務，乃立足於成熟的教育活動與應用，而且相對較長的文件與完整的句子，是其較具成效的重要因素。相對而言，簡短文字（如單句、子句、非完整文句、甚至是少數詞彙）的處理與應用，在創新的教育應用上，則較少被探討。

即時反饋系統（Instant Response Systems, IRS）、教室回應系統（Classroom Response Systems, CRS）、或者稱為按壓器（clickers）自 1960 年代起即在高等教育大班教室的上課中進行測試（Deal, 2007）。根據著名科技網站 CNET 的報導（Gilbert, 2005），美國中小學或大學單在 2004 一年中，就採購了近一百萬套按壓器。而光從兩個製造按壓器的廠商資料來看，這數字在十年內就累積了近九百萬套（Hoffman, 2012）。最近幾年，由於普及的行動設備以及成熟的雲端技術，使得教學現場的設備運用與資料蒐集更為簡易，IRS 在教室中更為普及（Bartsch & Murphy, 2011; Chen et al., 2013; Han, 2014; Morais, Barragués, & Guisasola, 2015）。國內 IRS 服務廠商如 Zuvio<sup>6</sup> 已在短時間內吸引國內校院的逐步使用。例如，在 2014 一年內台灣大學使用 Zuvio 的教師從 61 人增加到 263 人，從 68 門課增加到 384 門，從 2,037 位學生增加到 11,172 位（Lee & Shih, 2015）。

IRS 系統的運作方式，大致如下：教師與學生都有手持裝置，如手機或平板，透過 Web 網頁雲端平台，教師將題目播送給學生手上的行

---

2 <http://acl.ldc.upenn.edu/W/W03/index.html#W03-0200>, accessed on 2016/10/13.

3 [http://tm.itc.ntnu.edu.tw/CNLP/?q=NLP\\_TEA-2014](http://tm.itc.ntnu.edu.tw/CNLP/?q=NLP_TEA-2014), accessed on 2016/10/13.

4 <http://www.criterion.com.tw/>, accessed on 2016/10/13.

5 [http://turnitin.com/zh\\_tw/](http://turnitin.com/zh_tw/), accessed on 2016/10/13.

6 <http://www.zuvio.com.tw/>. Zuvio was developed since 2011, accessed on 2016/10/13.

動載具，每位學生的回應，立刻傳送回主機端並綜整呈現於教師的頁面上，再由教師決定是否公布結果，或是讓學生彼此的回應可以互相看到。一般的 IRS 都支援是非題、選擇題等電腦可輕易評估答案的題目，讓教師、學生立刻知道問項結果。如此，教師可以即時瞭解學生的學習狀態，從而採取措施幫助學生維持專注力，或是學生因為更瞭解其他同學的學習反應，增進了學生的學習投入狀況（Beatty & Gerace, 2009; Bartsch & Murphy, 2011）。

雖然有如此的效果，然而 IRS 的應用，尚有發揮的空間，且仍待進一步的探索與研究（Chien & Chang, 2015a）。以前述臺大的運用情形為例，教師出的題目當中，雖然有 54% 是選擇題，但仍有 20% 是開放性問題，用以提升學生的投注情形與反思機會（Lee & Shih, 2015）。這是因為不管是選擇題或是非題，學生仍有可能進行無意識的回應（隨機點選答案選項），但開放性問題則強迫學生以文字表達自己的想法，讓課堂的思考互動更深入。過去的研究甚至指出，選擇題在科學教育中會阻礙學生的高階（higher level）思考（Stanger-Hall, 2012），而建構式的回應（如自由的文字回應）被廣泛認為，可讓教師較易洞察學生思考的機會（Birenbaum & Tatsuoka, 1987）。其可能的應用情境如下：教師廣播一則開放性問題，學生據此輸入本身的想法來回應，這些短文回應（在有限的時間內，手持裝置較難輸入長篇大論）被即時歸類成不同的主題類別，讓教師可以快速得知教學期望與學生吸收的可能落差，以即時修正教學重點，甚至進一步據此將學生分組以便進行分組討論或辯論。此種情境，光以是非題來探究學生的學習並進行分組活動，即可成為有效的教學方式，若能以不受答案限制的開放性問題進行類似的活動，預期會有更佳的教學效果（Chien & Chang, 2015b）。

然而，目前還未有 IRS 提供開放性問答即時分析綜整批閱的功能。若能透過 IR / NLP 等技術的發展與應用，即時處理學生在課堂上的簡短回應，可以讓 IRS 的特性發揮得更好，進一步增進有意義的學習互動與有效的學習。惟，短文回應的自動處理，不見得比傳統長文處理來得容易。因為這些短文可能具有較高比例的雜訊（文句不完整、錯別字）、多國語言（中、英文夾雜）與用詞不一致等特性。

根據上述的趨勢與觀察，本文的目的，在探討 IRS 現況以及短文回應的特性與問題、回顧短文主題歸類自動處理的技術與應用情形、提出一套可行的搭配系統供 IRS 整合運用，並展示我們目前實作的結果範例，以做為後續討論、研究的參考。

## 貳、發展現況

目前 IRS 與短文處理，近幾年來都獲得相當多的關注，但匯聚兩者的研究與計畫稀少，因此底下分別介紹其發展現況。

### 一、即時反饋系統 IRS

上課中學生的投入參與 (engagement) 被視為是有效學習的關鍵指標之一 (Caldwell, 2007; Christopherson, 2011)。因此之故，教室回應系統 CRS 早在 1960 年代就被製作出來並進行測試 (Deal, 2007)。CRS 也被稱為 IRS、學生回應系統 (Student Response Systems, SRS)<sup>7</sup>、聽眾回應系統 (Audience Response Systems, ARS)、電子投票系統 (Electronic Voting Systems, EVS)，或是按壓器，且已創造出二億七千萬美元的市場 (Hoffman, 2012)。傳統的 IRS 包含一組 USB 無線接收器、教師端與學生端的按壓遙控器、以及相關軟體。其最常見的用法是教師將問題置於黑板或投影幕上，學生則在自己的遙控器上按壓適當的數字或像傳統電話按鍵那樣輸入英文字母。大家回應的結果則由教師透過其按壓器控制軟體做適當的呈現。傳統 IRS 支援數種問題規格，如：是非題、選擇題、填空題等。一些教育活動，如：小測驗、小問卷、票選、點名等都可以便利的進行，以刺激學生的參與投入 (Mayer et al., 2009)。

隨著智慧型手機的普及化，新一代的 IRS 幾乎都透過雲端網路環境 (App 或是 Web 網頁)，以免除課前需要發放按壓器的不便，並有利於紀錄與分析所有的 IRS 活動。表 1 比較了四種 IRS 的價格與功能。更多的 IRS 資訊可以在文獻 (Sanchez et al., 2012; Karakostas, Adam, Kioutsiouki, & Demetriadis, 2014) 或網頁<sup>8</sup> 上查到。其皆具備類似功能：使用的簡易性，並達成其增進學生課堂參與的目的。

在表 1 中開放性問題類型需要學生有意識的參與，並主動的回應，而非像是非題或選擇題那樣被限定答案，因此可以被動參與或隨意的亂按。所以多數的新型 IRS 都視開放性問題為必要功能。然而由於沒

7 <http://www.edweek.org/dd/articles/2009/06/17/04feedback.h02.html>, accessed on 2016/10/13.

8 iWonder 教學互動系統：[http://innovation.itmonth.org.tw/product\\_detail.php?pid=2015&uid=681&year=2015](http://innovation.itmonth.org.tw/product_detail.php?pid=2015&uid=681&year=2015)；網奕資訊 IRS 即時反饋系統：[http://www.habook.com.tw/eTeaching/products.aspx?BookNo=IRS\\_01](http://www.habook.com.tw/eTeaching/products.aspx?BookNo=IRS_01)，accessed on 2016/10/13.

有提供相關的自動化分析處理（如：主題偵測、語意歸類、關鍵詞擷取等文字比對、擷取功能），這些短文回應仍須教師一一瀏覽，以看出學生的想法，缺乏效率，而降低了教師的運用意願與頻率，也限制了 IRS 更大的發揮空間。另外，分組功能有助於學生投入同儕互助與辯論。然而，根據有限答案（是非題、選擇題）的回應所做的分組，仍有不足之處，從而降低了學生分組的教學益處。

除表 1 所列者之外，仍有更簡單的 IRS，像是 AnswerGarden<sup>9</sup>，有提供簡易的文字處理功能。AnswerGarde 允許 20 到 40 個字元的回應，然後以文字雲（word cloud）的方式呈現結果（此為其唯一的功能）。此種簡易的 IRS 仍可運用於教育活動上，如：語言學習<sup>10</sup>，但在本文中將不會視其為真正的 IRS，因為其缺乏表 1 中許多的功能。然而，不論哪一種 IRS，都需要教師運用創意突破教學白巢，以充分發揮 IRS 的功效（Beatty, Gerace, Leonard, & Dufresne, 2006; Deal, 2007）。

**表 1**  
**四種 IRS 的功能比較**

	CCR <sup>11</sup>	Zuvio <sup>12</sup>	SOCRATIVE <sup>13</sup>	傳統 IRS
約略價格（台幣）	免費	3000 元， 每班，每 學期	50 位學生以 上才開始計費	按壓器 1200 元
中文介面	✓	✓	×	✓
多國語文介面	中、英、法、 日、韓、阿拉 伯語	中、英	×	×
安裝費用	✓（WEB）	✓（WEB）	×	×
行動裝置的支援	✓	✓	✓	×

9 <http://answergarden.ch/>, accessed on 2016/10/22.

10 See examples at: <http://nikpeachey.blogspot.nl/2015/10/brainstorming-and-polling-with.html>, accessed on 2016/10/22.

11 Developed by Science Education Center, National Taiwan Normal University, since 2013.

12 Developed by the alumni of National Taiwan University since 2012 and has been commercialized.

13 A commercial product, for details, see: <http://www.socrative.com/>.

學生姓名顯示開關	✓	×	×	✓
題庫分享	✓	✓	✓	×
即時回應統計	✓	✓	✓	✓
回應匯出	✓	✓	✓	✓
與學校的學生資料介接	✓ (臺師大, 另可用 Facebook 帳號)	×	×	×
可用 PPT 做為問項	×	✓ (需額外安裝作業)	×	✓ (需額外安裝作業)
開放性問題類型	✓	✓	✓	×
事先將學生分組	✓	✓	×	✓
即時分組	✓	✓	✓	×
組內投票	✓	×	×	×
組內角色指派	✓	×	×	×
組內私訊	✓	×	×	×
同儕評鑑	×	✓	×	×

## 二、短文主題分析

近幾年來，短文處理在資訊檢索、自然語言處理、機器學習、文字探勘與數位學習等領域受到極大的關注 (Yang, Fan, Lai, Gao, & Wang, 2014; Vo & Ock, 2015; Zehner, Sälzer, & Goldhammer, 2015)。這是因為越來越多的短文出現在網路上，如：搜尋結果摘要、產品描述、客戶評價、討論串、聊天室、新聞短訊、微博、社會網路的訊息 (如：Facebook, Twitter)，以及教育評估的線上回應等。如同 Tang、Wang、Gao、Hu 與 Liu (2012) 指出，這些短文對傳統文字處理技術提出了新的挑戰：(1) 不夠充分的文字統計資訊；(2) 新詞或縮寫；(3) 同義異形詞；(4) 同形異義詞；(5) 拼錯字或錯別字；(6) 怪異的文法或不完整的文句。

挑戰 2 至 6 本來就會存在於任何文件中，只是短文的比例較高。挑戰 1、2、5、6 可能源自於手持裝置的輸入不便性 (時間、空間的侷促性)。挑戰 3 與 4 則多為挑戰 1 與 2 的延續問題。

IRS 系統裡的短文回應基本上也呈現出上述的問題。然而，卻極少有相關研究來處理這個問題。底下段落將回顧較為人知的短文（如推特訊息、新聞標題等）處理，說明其技術特性，看能否運用於 IRS 的短文回應。

從文獻中可摘要出兩種主要的策略來處理上述的挑戰：（1）以額外／外部資源增加短文的資訊量；（2）尋求更佳的模式對短文做語意的處理。

在策略 1 中，Banerjee、Ramanathan 與 Gupta（2007）對 Google 新聞標題作歸類時，提出運用與標題主題最相關的維基百科（Wikipedia）文章來提供額外的相關詞彙。Hu、Sun、Zhang 與 Chua（2009）在歸類 Google 搜尋結果摘要時，則先擷取其重要詞彙，再從詞網（WordNet）與維基百科來擴增語意相近的詞彙。Tang 等人（2012）則運用 Google 翻譯將短文翻譯成數種語言，以增進其文字量與特徵值，再透過矩陣分解來降低雜訊。在 Facebook 的使用者貼文以及推特訊息上，他們的擴增方法，比用原始短文更具成效。此結果，主要來自於 Google 幫忙翻譯後增加的前後文或是脈絡詞（contextual information），而解決了挑戰 1、2、3、4 的部分問題。此實驗顯示對某些語料翻譯成三、四種語言可得最佳效果，翻譯成更多語言則會引進太多雜訊而降低效果。

上述方法，均運用於英文短文。對中文而言，這些方法都需要實驗過，以驗證其成效。特別是中文語料（知識）資源遠比英文少，如中文詞網（Chinese WordNet，像 eHowNet<sup>14</sup>）資料量少於英文詞網（English WordNet），且中文維基百科條目（約 849,781）也大幅少於英文維基百科條目（約 5,019,983）<sup>15</sup>。

在策略二中，文件向量模式（Vector Space Model, VSM）是最經典傳統的方法（Salton, 1989）。其將語料中每份文件的重要詞彙（有主題意義的詞彙），都視為向量中的一個維度，而詞彙在文件中的出現次數（Term Frequency, TF）以及在整個語料中出現篇數的倒數（Inverse Document Frequency, IDF）的乘積（TF\*IDF），常做為該維度的權重。如此 N 篇文件的語料庫若共有 T 個詞彙，就形成一個 T\*N 的矩陣，其中每一行向量對應到每一篇文件，而每一列向量則對應到每一個重要詞

14 <http://ehownet.iis.sinica.edu.tw/index.php>, accessed on 2015/12/04.

15 [https://meta.wikimedia.org/wiki/List\\_of\\_Wikipedias](https://meta.wikimedia.org/wiki/List_of_Wikipedias), Statistics at 12:00, 29 November 2015 (UTC), accessed on 2016/10/30.

彙。依向量餘弦公式（或是其他相似度公式，如 Dice Coefficient），可計算任意兩文件的相似度（或是兩詞彙的相似度），從而運用各種歸類方法，如階層歸類（hierarchical clustering）、K-means 演算法等（Salton, 1989），將相似文件（或相似詞彙）歸類成數個主題群組。

VSM的問題，在於以詞彙為維度，若有不同詞彙卻語意相近時，因屬不同維度，也無法增加其相似度（即詞彙不匹配問題）。因此，在1990年左右，隱含語意索引法（Latent Semantic Indexing, LSI），或是稱做隱含語意分析法（Latent Semantic Analysis, LSA），被提出來（Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990）。其運用線性代數的奇異值分解（Singular Value Decomposition, SVD）方法，將 $T \times N$ 的矩陣降維（dimension reduction）成 $K \times K$ 的矩陣，其中 $K < T$ 且 $K < N$ 。亦即，語料中的文件篇數與重要詞彙個數，都被降為 $K$ 個維度，並讓此矩陣近似於整個語料，但個別文件的向量仍可從這個降維的矩陣算出近似值，然後以此算出任意兩篇文件的相似度，繼而將原來的文件進行歸類處理。此種降維的作法，讓語意相近的詞彙，被放在同一維度，解決了前述詞彙不匹配的缺點。例如，四個詞彙維度的文件（公車、巴士、醫院、病房），可被降維成二維的向量文件（ $1.3 \times \text{公車} + 0.4 \times \text{巴士} + 0.9 \times \text{醫院} + 0.8 \times \text{病房}$ ）。同理，主題相似的不同文件向量，也被濃縮到同一個主題的向量上，而這某種程度上解決了一詞多義的問題（因為可從該詞彙的前後文句脈絡詞，辨識其主題意義）。

之後發展出更精進的作法，機率式隱含語意索引法（probabilistic Latent Semantic Indexing, pLSI），其允許同一個文件，對應到多個主題（在每個主題上都有機率值）（Hofmann, 1999）。然而，pLSI讓新進的文件，難以融入既有的模型裡。因此，隱含狄立克雷分佈法（Latent Dirichlet Allocation, LDA）被發展出來（Blei, Ng, & Jordan, 2003）。

然而，LDA跟VSM一樣，將文件視為一群詞彙的集合，並沒有考慮到詞彙的順序。因此，對短文來說，可能還不是最佳的語意模型。為此，針對短文，研究者提出多種的語意模型，如：雙連主題模型（Bigram Topic Model）（Wallach, 2006）、非監督式主題分段模型（Unsupervised Topic Segmentation Model）（Jameel & Lam, 2013）、雙詞主題模型（Biterm Topic Model, BTM）（Cheng, Yan, Lan, & Guo, 2014）以及詞彙網路主題模型（Word Network Topic Model, WNTM）（Zuo, Zhao, & Xu, 2015）。其中BTM的延伸作法被運用在東森新聞中文標題的歸類上，且有良好的成效（Chen, 2015）。

近年來，密西根州立大學（Michigan State University）曾將文字

自動分析的方法，運用於書寫的形成性評量上 (Urban-Lurain et al., 2013)。他們以 IBM-SPSS Modeler 工具中的 K-means 歸類法，將學生的文字回應分成數群，在一個工作天內處理完並做成報告，提供老師於下次上課前理解學生狀況。如圖 1 所示，第一列為老師給的問題，其下為 174 位學生回應的歸類結果 (包含其用詞或概念)、其人工撰寫的類別描述、類別中代表性的學生回應範例 (距離類別向量中心點最近的回應向量)、以及一個網狀圖以顯示詞彙間的關係 (兩個詞彙之間的連線代表學生傾向會同時寫出這兩個詞彙)。學生的回應，在報告中大多被歸為 3 到 5 類。其針對教師的焦點團體訪談顯示，此回饋報告有助於洞察學生「為何」學習有困難，而不像傳統選擇題那樣只能知道「哪裡」有困難。然而針對學生文字回應的分析並產生這樣的報告，需花費數個工作小時，對 IRS 系統而言，這是不可行的 (infeasible)。不過這篇文章，是我們找到「將學生的短文回應文字歸類整理起來，供老師快速瞭解學生學習狀況」，唯一最直接相關的文獻。其中的一些分析，如：類別主題描述、摘要學生回應範例、學生回應的關聯詞視覺化，雖然在其他類型的文件 (或詞彙) 歸類中有見過個別的運用，如：Tseng (2010b)、Tseng 與 Tsay (2013)、Tseng、

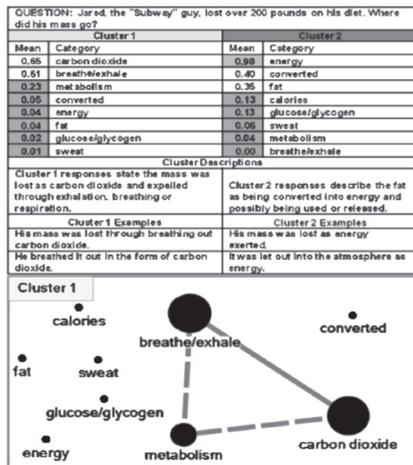


圖 1 密西根州立大學學生回應文字的自動分析報告範例

Urban-Lurain, M., Prevost, L., Haudek, K. C., Henry, E. N., Berry, M., & Merrill, J. E. (2013). Using computerized lexical analysis of student writing to support Just-in-Time teaching in large enrollment STEM courses. Paper presented at the Frontiers in Education Conference, 2013 IEEE, Oklahoma City, OK.

Chang、Rundgren Chang 與 Rundgren (2010) 等文所述，但此篇文章啟發了我們可以將這些歸類與分析一起運用在 IRS 系統上的可行性。

總結而言，短文自動處理近幾年來受到學術界極大的關注，但多聚焦在 Twitter、Weibo、Facebook 貼文或新聞標題等文件的主题擷取、歸類、偵測上。我們尚未發現有直接應用於 IRS 短文回應上的案例。這方面的研究，有待持續投入，以進一步提升 IRS 的應用成效。

## 參、短文主題歸類系統

本研究的目標為提供一套適用於 IRS 的文字處理系統，至少能夠做到關鍵詞彙的擷取、主题的分析與組織、與文件的歸類分群等功能，讓老師可以快速掌握學生的自由回應，並據此進行合適的教學活動，充分利用 IRS 的特性，來提升教學成效。為此，研究中選擇一套既有的 IRS：CloudClassRoom (CCR, <http://ccr.tw/>) (Chien & Chang, 2015a) 與其整合，來試驗此系統的功能。然而，在發展此系統時，我們也保持其模組化與可移植的特性，使其亦可易於跟其他的 IRS 整合。

圖 2 是我們提出的短文處理流程分析模組 (Short Text Analysis Module, STAM)，其中的方塊代表處理單元，圓柱代表所需的語言知識、語料或技術選項等。圖 2 的上半部流程主要在處理回應文字的詞彙細節，而這跟語言知識與資源有關；下半部的流程則處理文意或語意，其技術與哪種語言較不相關。此種流程架構是必要的，因為不同語言（如中文或英文）、不同性質的課程，在處理過程中，會運用不同的選項。在發展初期，這些選項由教師或文字處理專家來設定、調教；試用過一段時間累積經驗後，期望可以由系統自行調整。例如，在社會學科進行語意的歸類，對學生的回應做分群時，不同的數字，可能只要轉成相同的替代符號即可，但在數學、工程等學科，學生回應中的數字必須保留不更動。另外，一般情形下，英文詞彙需先取詞幹做型態分析 (morphological analysis) 等處理，使名詞或動詞的單複數、時態，不影響語意；但在英文課中，這些詞彙型態分析可能要關閉不做才符合需求。這些考慮可讓 STAM 達到最佳化，但隨著應用的課程與情境越來越多，需要考慮更多選項，才能適用於各種教學活動。事實上，CCR 目前已有至少 486 位教師註冊、建立 2529 間虛擬教室，以及超過一萬筆各種型態（包含是非、選擇、填充、開放性問題）的問答活動。要完全滿足各種應用情境，需要多年的調校與相關語料資源的融合，才能使 STAM 具備完整的實用性。

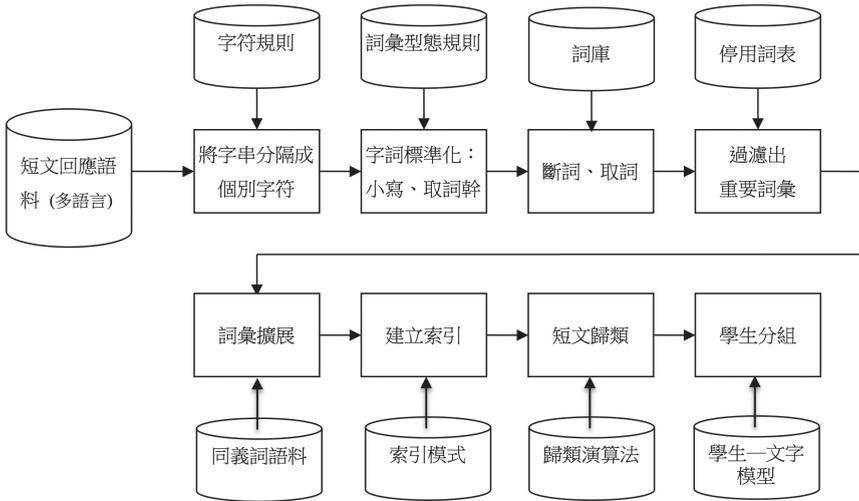


圖 2 應用於 IRS 的短文分析流程模組

## 肆、實例展示

目前初步的實作系統已整合入 CCR，並可進行短文回應語意分群的簡單應用。如圖 3 所示，針對研究所「資訊計量學」課程，教師提出問題，透過測試中的語意自動分類功能，將學生回應歸類成 3 組，以快速瞭解 6 位學生的回應情形。



圖 3 CCR 系統測試中的短文回應自動歸類功能—以資訊計量學的題目為範例

從圖 3 可看出，系統的歸類功能，有其合理之處，亦有可商榷之處。若將學生編號 9445 與 11737 的回應歸在組別 1，組別 3 剩下兩個歸成 2 類，似乎更為理想（此為主觀看法，因為歸類沒有標準答案）。

底下展示另一項實例，以更瞭解學生回應的短文特性，並應用不同技術來實作語意歸類，以瞭解各種方法的特性與優劣，以及後續需要注意、加強之處。

圖 4 為授課中科普教師顯示圖片予學生後，透過 CCR 提問：「假若你是一位海洋生物研究者，有人拍攝到一隻奇特的生物，希望詢問你的意見並請你進行初步的推測。請問你想先問什麼問題？會推測其有何特性？為什麼？」，表 2 則顯示學生以手持設備回應的短文。

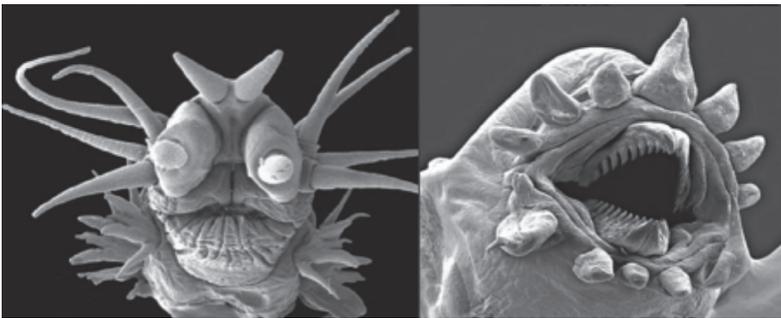


圖 4 科普教學中用以激發學生回應的圖片

表 2

學生針對圖 4 的提問，透過 CCR 所做的短文回應

ID	底下為 29 位學生的回應
1	1. 發現地點 2. 推論有毒 3. 外星生物 地球沒有
3	1 好吃嗎。2 肉食性。3 牙齒很尖。4 深海魚。5 因為很醜 ==
4	What's its life cycle? I guess that it's a meat-eater. Maybe it's a parasite. Because it has a fixation structure.
5	他們有毒嗎？
6	他們看得到嗎？他們的食物可能是什麼？
9	他是生活在何種海域？！深度？！環境？！
11	他的體型大小
12	住在很深的海裡吧！眼睛很凸，牙齒很尖應該是食肉動物！

15	呵呵呵呵呵
17	問題：在哪個海域發現的呢？特性：吃腐肉，極可能是古老的活化石。 原因：長相特別、牙齒尖銳。生物：恐龍？！或其相關生物
18	問題：水深位置大約在何處。推測：疑似刺絲胞動物門，有攻擊力，有尖銳的外型
21	好奇怪
22	它有肛門嗎？
23	對光源有無反應
24	很像大英雄天團的噴火龍
25	我覺得它是人類的祖先。因為他有眼睛 有嘴巴 有牙齒。牠以細菌為生， 牠屬於夜行性動物，睡眠時間為 12 小時 也就是半天，是個奇特的生物!!! 我想要 usb
27	海洋生物
28	深海生物，無視覺
29	爸爸
31	牠生活於海洋表層還深層？授獵能力較強，因為牙齒以犬齒較多可進行撕裂，海洋深層消費者
32	發現的環境包含深度 身體外表特徵，實行生物分類 它的捕食習性
33	眼睛會感光嗎？肉食性動物，牙齒看起來很尖，深海的未知生物，因為 看不出來是什麼種類的生物
34	神奇寶貝
35	肉食性的魚，在很深暗海
36	觸角是類似珊瑚的觸角嗎，應該是住在深海裡的變種生物吧。
37	跟我同學好像
38	身體構造有那些特徵
39	這個生物是不是小小隻的？可能是吃浮遊生物的，深海的生物，因為有觸手
43	這種生物有攻擊性嗎？應該住在深海？有照明的能力吧！這會不會是鯊魚和燈籠魚的合體

如表 2 所示，學生的短文回應有下列特徵：（1）無意義的標點符號（如文件編號 ID3）；（2）在中文課堂中用英文回應（如 ID4）；（3）無理的回應（如 ID15、24、29）；（4）非常短的文字（如 ID5、11、27）；以及（5）非主題文字（如 ID25 最後的部分）。

理想的作法，可以將特徵 1 在分隔字符（tokenization）階段去除。

特徵 2 可利用辭典資源（如：雙語辭典或是多語詞網，像 BabelNet<sup>16</sup>）以簡易的字對字、詞對詞的翻譯，或是用開源碼的客制化機器翻譯技術（Chuang & Tseng, 2008; Tseng et al., 2011）將其翻譯成中文。特徵 4 可透過同義詞辭典或是多語詞網來增加其語意資訊。但即使有中研院的 eHowNet 資源，也無法保證其能涵蓋各類科目的知識範圍，以進行同義詞或是上位詞的擴展。理想上，若用這些語言知識、資源與工具做完前處理後，短文回應的語意處理難度已可降低，接下來即可進行短文回應的歸類處理。

在蒐集足夠的語言資源、工具前，本研究嘗試直接處理短文、同時運用三種既有方法進行比較，並對其成效做初步的探索，以瞭解後續資源需求、更佳技術的殷切程度。此三種方法分別為：

1. 階層聚合歸類法（Hierarchical Agglomerative Clustering, HAC）：  
以混和的詞彙索引方法，亦即：以內建詞庫來斷詞，並用自動方法來擷取新詞或關鍵詞（Tseng, 1998, 2002; Tseng et al., 2010），且移除停用詞與標點符號，依此將每篇短文轉換成文件向量，再進行相似度運算後，以 HAC 進行文件歸類。此方法已實作於 CATAR 工具（Tseng, 2010a; Tseng & Tsay, 2013），且經過數年的測試與運用，算是相當成熟的方法。其結果如圖 5。
2. 隱含語意分析法（Latent Semantic Analysis, LSA）：  
根據開源碼 jieba<sup>17</sup> 透過其內建詞庫及條件隨機域（Conditional Random Field, CRF）技術進行斷詞與新詞處理，在沒有過濾停用詞與標點符號的情況下（因為在短文中，這兩種也可以成為辨識相似說法習慣的特徵），再根據免費的主題模型工具 gensim<sup>18</sup> 進行文件歸類。其結果如圖 6。事實上，圖 3 右邊的結果，即是運用此方法得出的結果。
3. 隱含狄立克雷分佈法（Latent Dirichlet Allocation, LDA）：  
與方法 2 相同，也是運用 jieba 與 gensim，但有移除停用詞以及標點符號。其結果如圖 7。

由於可用來試驗的技術與資源，有各種選項，排列組合極多，本文僅以上述三種方法，初步的探究短文歸類在協助教師組織短文、分析資訊，而不用一一瀏覽全部短文的效用。

---

16 <http://babelnet.org/>, accessed on 2015/12/03.

17 <https://github.com/fxsjy/jieba>, accessed on 2015/12/03.

18 <https://radimrehurek.com/gensim/>, accessed on 2015/12/03.

- 7 Docs. : 0.27 ( 生物 :14.40, 深海 :1.90)
  - 4 Docs. : 0.96 ( 深海 :5.76, 生物 :5.33)
    - 3 Docs. : 1.00 ( 深海 :3.5496, 生物 :2.1909)
      - 43 : 這種生物有攻擊性嗎？應該住在深海？有照明的能力吧！這會不會是鯊魚和燈籠魚的合體
      - 36 : 觸角是類似珊瑚的觸角嗎，應該是住在深海裡的變種生物吧。
      - 28 : 深海生物，無視覺
    - 39 : 這個生物是不是小小隻的？可能是吃浮遊生物的，深海的生物，因為有觸手
  - 3 Docs. : 0.86 ( 發現 :6.0000, 生物 :2.92)
    - 2 Docs. : 1.00 ( 發現 :3.11, 分類 :1.35, 有毒 :1.35, 地球 :1.35, 推論 :1.35)
      - 1 : ? 發現地點？推論有毒？外星生物 地球沒有
    - 32 : 發現的環境包含深度 身體外表特徵，實行生物分類 它的捕食習性
    - 17 : 問題：在哪個海域發現的呢？特性：吃腐肉，極可能是古老的活化石。原因：長相特別、牙齒尖銳。生物：恐...
- 3 Docs. : 0.73 ( 肉食性 :6.00, 牙齒 :1.27, 深海 :0.90)
  - 2 Docs. : 0.86 ( 肉食性 :3.11, 肉食性的魚 :1.35, 好吃 :1.35, 牙齒 :0.20, 深海 :0.06)
    - 35 : 肉食性的魚，在很深暗海
    - 3 :1 好吃嗎。2 肉食性。3 牙齒很尖。4 深海魚。5 因為很醜 ==
  - 33 : 眼睛會感光嗎？肉食性動物，牙齒看起來很尖，深海的未知生物，因為看不出來是什麼種類的生物
- 3 Docs. : 0.70 ( 動物 :4.92, 牙齒 :1.27)
  - 2 Docs. : 0.90 ( 動物 :2.55, 牙齒 :2.15, 奇特 :1.35, 細菌 :1.35, 食肉 :1.35)
    - 12 : 住在很深的海裡吧！眼睛很凸，牙齒很尖應該是食肉動物！
    - 25 : 我覺得它是人類的祖先。因為他有眼睛有嘴巴 有牙齒。牠以細菌為生，牠屬於夜行性動物，睡眠時間為 12 小時 ...
  - 18 : 問題：水深位置大約在何處。推測：疑似刺絲胞動物門，有攻擊力，有尖銳的外型
- 1 Docs. : 4 : What's its life cycle? I guess that it's a meat-eater. Maybe it's a parasite. Because it has a fixation structure.
- 1 Docs. : 5 : 他們有毒嗎？
- 1 Docs. : 6 : 他們的食物可能是什麼？
- 1 Docs. : 9 : 他是生活在何種海域 ?? 深度 ?? 環境 ??
- 1 Docs. : 11 : 他的體型大小
- 1 Docs. : 22 : 它有肛門嗎？
- 1 Docs. : 27 : 海洋生物
- 1 Docs. : 23 : 對光源有無反應
- 1 Docs. : 31 : 牠生活於海洋表層還深層？授獵能力較強，因為牙齒以大齒較多可進行撕裂，海洋深層消費者
- 1 Docs. : 38 : 身體構造有那些特徵
- 1 Docs. : 21 : 好奇怪
- 1 Docs. : 37 : 跟我同學好像
- 1 Docs. : 15 : 呵呵呵呵呵
- 1 Docs. : 24 : 很像大英雄天團的噴火龍
- 1 Docs. : 29 : 爸爸
- 1 Docs. : 34 : 神奇寶貝

圖 5 階層聚合歸類方法 (HAC) 的歸類結果

ID	詞彙關聯圖	文件 ID : 回應文字
1		<p>3 : 1 好吃嗎。2 肉食性。3 牙齒很尖。4 深海魚。5 因為很醜 ==</p> <p>17 : 問題 : 在哪個海域發現的呢? 特性 : 吃腐肉, 極可能是古老的活化石。原因 : 長相特別、牙齒尖銳。生物 : 恐龍?! 或其相關生物</p> <p>18 : 問題 : 水深位置大約在何處。推測 : 疑似刺絲胞動物門, 有攻擊力, 有尖銳的外型</p> <p>31 : 牠生活於海洋表層還深層? 授獵能力較強, 因為牙齒以犬齒較多可進行撕裂, 海洋深層消費者</p> <p>33 : 眼睛會感光嗎? 肉食性動物, 牙齒看起來很尖, 深海的未知生物, 因為看不出來是什麼種類的生物</p> <p>36 : 觸角是類似珊瑚的觸角嗎, 應該是住在深海裡的變種生物吧。</p> <p>39 : 這個生物是不是小小隻的? 可能是吃浮遊生物的, 深海的生物, 因為有觸手</p> <p>43 : 這種生物有攻擊性嗎? 應該住在深海? 有照明的能力吧! 這會不會是鯊魚和燈籠魚的合體</p>
2		<p>5 : 他們有毒嗎?</p> <p>6 : 他們看得到嗎? 他們的食物可能是什麼?</p>
3		<p>11 : 他的體型大小</p> <p>25 : 我覺得它是人類的祖先。因為他有眼睛 有嘴巴 有牙齒。牠以細菌為生, 牠屬於夜行性動物, 睡眠時間為 12 小時 也就是半天, 是個奇特的生物 !!! 我想要 usb</p> <p>32 : 發現的環境包含深度 身體外表特徵, 實行生物分類 它的捕食習性</p> <p>37 : 跟我同學好像</p> <p>38 : 身體構造有那些特徵</p>

4		<p>1：1. 發現地點 2. 推論有毒 3. 外星生物 地球沒有</p> <p>4：What's its life cycle? I guess that it's a meat-eater. Maybe it's a parasite. Because it has a fixation structure.</p> <p>9：他是生活在何種海域？！深度？！環境？！</p> <p>22：它有肛門嗎？</p> <p>23：對光源有無反應</p>
5		<p>12：住在很深的海裡吧！眼睛很凸，牙齒很尖應該是食肉動物！</p> <p>24：很像大英雄天團的噴火龍</p> <p>28：深海生物，無視覺</p> <p>35：肉食性的魚，在很深暗海</p>
N/A		<p>15：呵呵呵呵呵</p> <p>21：好奇怪</p> <p>27：海洋生物</p> <p>29：爸爸</p> <p>34：神奇寶貝</p>

圖 6 隱含語意分析方法 (LSA) 的歸類結果，其中詞彙關聯圖是用 Tseng (2002) 的方法做出來的

ID	詞彙關聯圖	文件 ID：回應文字
1		<p>1：1. 發現地點 2. 推論有毒 3. 外星生物 地球沒有</p> <p>5：他們有毒嗎？</p> <p>6：他們看得到嗎？他們的食物可能是什麼？</p> <p>21：好奇怪</p> <p>22：它有肛門嗎？</p>
2		<p>31：牠生活於海洋表層還深層？授獵能力較強，因為牙齒以犬齒較多可進行撕裂，海洋深層消費者</p> <p>39：這個生物是不是小小隻的？可能是吃浮遊生物的，深海的生物，因為有觸手</p> <p>37：跟我同學好像</p> <p>27：海洋生物</p>

<p>3</p>		<p>12：住在很深的海裡吧！眼睛很凸，牙齒很尖應該是食肉動物！</p> <p>33：眼睛會感光嗎？肉食性動物，牙齒看起來很尖，深海的未知生物，因為看不出來是什麼種類的生物</p> <p>43：這種生物有攻擊性嗎？應該住在深海？有照明的能力吧！這會不會是鯊魚和燈籠魚的合體</p> <p>36：觸角是類似珊瑚的觸角嗎，應該是住在深海裡的變種生物吧。</p> <p>9：他是生活在何種海域？！深度？！環境？！</p> <p>15：呵呵呵呵呵</p>
<p>4</p>		<p>18：問題：水深位置大約在何處。推測：疑似刺絲胞動物門，有攻擊力，有尖銳的外型</p> <p>28：深海生物，無視覺</p> <p>3：1 好吃嗎。2 肉食性。3 牙齒很尖。4 深海魚。5 因為很醜 ==</p> <p>23：對光源有無反應 32：發現的環境包含深度 身體外表特徵，實行生物分類 它的捕食習性</p> <p>38：身體構造有那些特徵</p> <p>4：What's its life cycle?I guess that it's a meat-eater. Maybe it's a parasite.Because it has a fixation structure.24：很像大英雄天團的噴火龍</p> <p>29：爸爸</p>
<p>5</p>		<p>17：問題：在哪個海域發現的呢？特性：吃腐肉，極可能是古老的活化石。原因：長相特別、牙齒尖銳。生物：恐龍？！或其相關生物</p> <p>35：肉食性的魚，在很深暗海</p> <p>11：他的體型大小</p> <p>25：我覺得它是人類的祖先。因為他有眼睛 有嘴巴 有牙齒。牠以細菌為生，牠屬於夜行性動物，睡眠時間為 12 小時 也就是半天，是個奇特的生物 !!! 我想要 usb</p> <p>34：神奇寶貝</p>

圖 7 隱含狄立克雷分佈方法 (LDA) 的歸類結果

圖 5 顯示，HAC 歸類出 3 個多篇文件的群組，以及 16 個單篇文件的群組。仔細端詳，其結果有一定程度的合理度。例如，第一類 7 篇文件有兩個小類，其第一小類的短文都有：「深海」、「生物」的詞彙，而第二小類則都有：「發現」、「生物」這兩個詞彙。因此，

第一類主題為「生物」，再以「深海」及「發現」地區來區分小類。第二個類別全部 3 篇短文均出現「肉食性」一詞，其中 2 篇出現「深海」、「牙齒」。因此，第二類的主題，可視為是「肉食性」生物，雖然也包含「深海」、「牙齒」的概念。第三個類別有 3 篇出現「動物」，2 篇出現「牙齒」。因此，第三類主題主要為「動物」，其次也包含「牙齒」的概念。其實這三類：「生物」、「肉食性」、「動物」在概念上幾乎是相同的主題，嚴格區分才有差別，而且又有「深海」、「牙齒」等詞彙交錯在不同類別中，似乎很難分類區別。但 HAC 能夠在這麼聚焦的小主題上，透過圖 5 中的類別標題詞彙，分出三個細微類別，已實屬不易。即便如此，一些文件的歸類也有不適當之處，如：文件 ID9、27 與 31 等，沒有與其他相似文件歸類在一起，則是因為詞彙不匹配問題，如「海域、深度」及「海洋、深層」跟「深海」不匹配，再加上此處的 HAC 使用較嚴格的歸類條件（完全連結相似度，complete-link similarity），以致於這些文件沒有跟其他文件歸類在一起。類似的問題，也出現在「光源」（ID23）、「感光」（ID33），或「食物」（ID6）與「肉食性」（ID3，35）。

為了讓語意相似的文件歸類在一起，解決詞彙不匹配問題，使用 LSA 或 LDA 似乎是合理的選擇。然而從圖 6 及圖 7 的結果顯示，若不能對 LSA 與 LDA 做更深入的應用（如運用更多語料資源做好上述的前處理動作），目前直接套用的方式，就不一定可以得出較佳的結果。例如圖 6 中，雖然其第 3 類可用此生物的「身體特徵」做為主題，並且也將多篇無理的（nonsense）回應區別出來到第 6 類（N/A）上；但在第 2 類中，此生物是否有毒跟其食物關聯度不高，而第 4 類中的 5 篇短文則似乎都不相關。圖 7 的 LDA 結果，也有不合理之處，總有些無理的回應混雜其中，應是短文回應的語料不足（僅 29 篇）以及文件太短共現詞（co-occurrence terms）不足，而無法完全發揮 LDA 的優勢所致。

總結而言，上述方法，都有協助組織短文、分析資訊的功用。但影響短文語意歸類成效的因素，不只有語意模型、歸類方法，還有詞彙索引，例如，是否用單連詞（1-gram）、雙連詞（2-gram）、詞彙擴展（如「海域、深度」及「海洋、深層」跟「深海」，「光源」與「感光」）等因素。此外，這些因素，也可能隨题目的特性、所屬科目之不同，而有不同的重要性。這些都需要更多的測試、觀察、應用、回饋與調校等費時的循環步驟，才有機會深入瞭解各種選項的特性，進而針對各種提問與回應，做出最佳的成效。

## 伍、結語

本文介紹了教育領域最近發展的 IRS 現況、文字歸類與分析應用在 IRS 的情形與重要性，說明 IRS 中短文回應的特性，進而提出一套可行的短文主題歸類模組，最後以真實案例呈現結果，並討論各種因素，比較三種初步歸類方法的成效。

此處短文主題自動歸類的成效，如同多數的 IR / NLP 任務一樣，有各種影響因素（曾元顯，2002；曾元顯、王峻禧，2007）。雖本研究尚未具全面性的實驗結果（可以組合的實驗參數眾多、需要的人工類別標記費時），前一節的分析與比較，已印證了在理論上較好的歸類方法（如 LSA 或 LDA），不見得會有較佳的結果，還需其他因素或處理流程的配合。

後續的研究，將蒐集各種語言資源，試驗各種前處理的方案，繼續尋求最佳的短文主題歸類方法，並進行較大規模實驗的成效比較。最重要的，就是將主題歸類最佳方案整合到 IRS 後，進行實際的教學測試，並評估其應用成效。

上述兩項工作均需時間累積更多資源、精進技術與經驗與進行教學應用測試。若有時間、資源往此方向繼續投入，相信未來可以提供真正有用的短文主題分析模組以及最佳實施經驗，供學術界與產業界有效的應用。

本文提出整合在 CCR 中的 STAM 模組，理想上可應用在：社會學科、數學、工程、英文、科學教育等多種學科，文中我們描述了其應用的可能性、注意事項以及範例。然而，跟所有的教學科技一樣，IRS 的開放性問題，也需要老師透過學習、體驗、累積經驗後，才會獲得良好的教學效果。本研究之系統，也將根據紀錄檔中的資料以及後續老師回饋的意見，加以改善。目標是使其盡量適用於高中以上各種學科的創意教學上。

## 誌謝

本研究感謝教育部「邁向頂尖大學計畫」、國立臺灣師範大學「華語文與科技研究中心」、科學教育中心，以及科技部專題研究計畫：MOST 105-2221-E-003-020-MY2 之補助。

## 參考文獻

- 曾元顯 (2002)。文件主題自動分類成效因素探討。中國圖書館學會會報, 68, 62-83。
- 曾元顯、王峻禧 (2007)。分類不一致之自動偵測—以農資中心資料為例。圖書館學與資訊科學, 33 (2), 19-32。
- Banerjee, S., Ramanathan, K., & Gupta, A. (2007). *Clustering short texts using wikipedia*. Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, 787-788. doi: 10.1145/1277741.1277909
- Bartsch, R. A., & Murphy, W. (2011). Examining the Effects of an Electronic Classroom Response System on Student Engagement and Performance. *Journal of Educational Computing Research*, 44(1), 25-33. doi:10.2190/EC.44.1.b
- Beatty, I., & Gerace, W. (2009). Technology-Enhanced Formative Assessment: A Research-Based Pedagogy for Teaching Science with Classroom Response Technology. *Journal of Science Education and Technology*, 18(2), 146-162. doi:10.1007/s10956-008-9140-4
- Beatty, I. D., Gerace, W. J., Leonard, W. J., & Dufresne, R. J. (2006). Designing effective questions for classroom response system teaching. *American Journal of Physics*, 74(1), 31-39. doi:http://dx.doi.org/10.1119/1.2121753
- Birenbaum, M., & Tatsuoka, K. K. (1987). Open-Ended Versus Multiple-Choice Response Formats—It Does Make a Difference for Diagnostic Purposes. *Applied Psychological Measurement*, 11(4), 385-395. doi:10.1177/014662168701100404
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Caldwell, J. E. (2007). Clickers in the large classroom: current research and best-practice tips. *CBE-Life Science Education*, 6(1), 9-20. doi:10.1187/cbe.06-12-0205
- Cardoso, W. (2010). Clickers in foreign language teaching: a case study. *Contact: Teachers of English as a Second Language of Ontario*, 36(2), 36-55.
- Chen, G. B. (2015). *Word Co-occurrence Augmented Topic Model in Short*

- Text*. (Unpublished master's thesis), National Cheng Kung University, Tainan, Taiwan.
- Chen, T. L., Lin, Y. F., Liu, Y. L., Yueh, H. P., Sheen, H. J., & Lin, W. J. (2013). Integrating Instant Response System (IRS) as an In-Class Assessment Tool into Undergraduate Chemistry Learning Experience: Student Perceptions and Performance. In M. H. Chiu, H. L. Tuan, H. K. Wu, J. W. Lin, & C. C. Chou (Eds.), *Chemistry Education and Sustainability in the Global Age* (pp. 267-275). Netherlands: Springer.
- Cheng, X. Q., Yan, X. H., Lan, Y. Y., & Guo, J. F. (2014). BTM: Topic Modeling over Short Texts. *Ieee Transactions on Knowledge and Data Engineering*, 26(12), 2928-2941. doi:10.1109/tkde.2014.2313872
- Chien, Y. T., & Chang, C. Y. (2015a). *Providing students with an alternative way to interact with the teacher in the silent classroom: Teaching with the CloudClassRoom technology*. Poster session presented at the Inaugural Asian Conference on Education & International Development (ACEID), Osaka, Japan.
- Chien, Y. T., & Chang, C. Y. (2015b). Supporting socio-scientific argumentation in the classroom through automatic group formation based on students' real-time responses. In M. S. Khine (Ed.), *Science education in East Asia: Pedagogical innovations and research-informed practices* (pp. 549-563). Springer: International Publishing.
- Christopherson, K. M. (2011). Hardware or wetware: what are the possible interactions of pedagogy and technology in the classroom? *Teaching of Psychology*, 38, 288-292. doi:10.1177/0098628311421332
- Chuang, Z. J., & Tseng, Y. H. (2008). NTCIR-7 Experiments in Patent Translation based on Open Source Statistical Machine Translation Tools. Paper presented at the *Seventh NTCIR Workshop on Evaluation of Information Access Technologies: Information Retrieval, Question Answering, and Cross-Lingual Information Access*, Tokyo, Japan.
- Deal, A. (2007). *A Teaching with Technology White Paper: Classroom Response Systems*. Retrieved from [https://www.cmu.edu/teaching/technology/whitepapers/ClassroomResponse\\_Nov07.pdf](https://www.cmu.edu/teaching/technology/whitepapers/ClassroomResponse_Nov07.pdf)
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391-407. doi:10.1002/

- (SICI)1097-4571(199009)41:6<391::AID-ASII>3.0.CO;2-9
- Gilbert, A. (2005). *New for back-to-school: 'Clickers'*. Retrieved from <http://www.cnet.com/news/new-for-back-to-school-clickers/>
- Han, J. H. (2014). Closing the Missing Links and Opening the Relationships among the Factors: A Literature Review on the Use of Clicker Technology Using the 3P Model. *Journal of Educational Technology & Society*, 17(4), 150-168.
- Hearst, M. A. (2015). Can Natural Language Processing Become Natural Language Coaching? Paper presented at *The 53rd Annual Meeting of the Association of Computational Linguistics*, Beijing, China.
- Hoffman, J. (2012). *Speak Up? Raise Your Hand? That May No Longer Be Necessary*. Retrieved from [http://www.nytimes.com/2012/03/31/us/clickers-offer-instant-interactions-in-more-venues.html?\\_r=0](http://www.nytimes.com/2012/03/31/us/clickers-offer-instant-interactions-in-more-venues.html?_r=0)
- Hofmann, T. (1999). *Probabilistic latent semantic indexing*. Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, 50-57. doi: 10.1145/312624.312649 Retrieved from <http://dl.acm.org/citation.cfm?doid=312624.312649>
- Hu, X., Sun, N., Zhang, C., & Chua, T. S. (2009). *Exploiting internal and external semantics for the clustering of short texts using world knowledge*. Proceedings of the 18th ACM conference on Information and knowledge management, 919-928. doi: 10.1145/1645953.1646071 Retrieved from <http://dl.acm.org/citation.cfm?doid=1645953.1646071>
- Jameel, S., & Lam, W. (2013). *An unsupervised topic segmentation model incorporating word order*. Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, 203-212. doi: 10.1145/2484028.2484062
- Karakostas, A., Adam, D., Kioutsiouki, D., & Demetriadis, S. (2014). A pilot study of QuizIt: The new android classroom response system. Paper presented at the *2014 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL)*, Thessaloniki, Greece.
- Lee, J. W. S., & Shih, M. L. (2015). Teaching Practices for the Student Response System at National Taiwan University. *International Journal of Automation and Smart Technology*, 5(3), 145-150. doi:10.5875/

ausmt.v5i3.862

- Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., ... Zhang, H. J. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, 34(1), 51-57. doi:10.1016/j.cedpsych.2008.04.002
- Morais, A., Barragués, J. I., & Guisasola, J. (2015). Using a Classroom Response System for promoting interaction to teaching Mathematics to large groups of undergraduate students. *Journal of Computers in Mathematics and Science Teaching*, 34(3), 249-271.
- Salton, G. (1989). *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*. Reading, MA: Addison-Wesley.
- Sanchez, C. G., Castro, F., Gomez, J. I., Tenllado, C., Chaver, D., & Lopez-Orozco, J. A. (2012). *OpenIRS-UCM: an open-source multi-platform for interactive response systems*. Proceedings of the 17th ACM annual conference on Innovation and technology in computer science education, 232-237. doi: 10.1145/2325296.2325352 Retrieved from <http://dl.acm.org/citation.cfm?doid=2325296.2325352>
- Siau, K., Hong, S., & Nah, F. F. H. (2006). Use of a classroom response system to enhance classroom interactivity. *IEEE Transactions on Education*, 49(3), 398-403. doi:10.1109/TE.2006.879802
- Stanger-Hall, K. F. (2012). Multiple-Choice Exams: An Obstacle for Higher-Level Thinking in Introductory Science Classes. *CBE Life Sciences Education*, 11(3), 294-306. doi:10.1187/cbe.11-11-0100
- Tang, J. L., Wang, X. F., Gao, H. J., Hu, X., & Liu, H. (2012). Enriching short text representation in microblog for clustering. *Frontiers of Computer Science*, 6(1), 88-101. doi:10.1007/s11704-011-1167-7
- Tseng, Y. H. (1998). *Multilingual Keyword Extraction for Term Suggestion*. Proceedings of the 21st International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '98, 377-378. doi: 10.1145/290941.291066
- Tseng, Y. H. (2002). Automatic Thesaurus Generation for Chinese Documents. *Journal of the American Society for Information Science and Technology*, 53(13), 1130-1138.

- Tseng, Y. H. (2010a). *Content Analysis Toolkit for Academic Research (CATAR)*. Retrieved from <http://web.ntnu.edu.tw/~samtseng/CATAR/>
- Tseng, Y. H. (2010b). Generic Title Labeling for Clustered Documents. *Expert Systems With Applications*, 37(3), 2247-2254.
- Tseng, Y. H., Chang, C. Y., Rundgren Chang, S. N., & Rundgren, C. J. (2010). Mining Concept Maps from News Stories for Measuring Civic Scientific Literacy in Media. *Computers & Education*, 55(1), 165-177. doi:10.1016/j.compedu.2010.01.002
- Tseng, Y. H., Liu, C. L., Tsai, C. C., Wang, J. P., Chuang, Y. H., & Jeng, J. (2011). Statistical Approaches to Patent Translation for PatentMT-Experiments with Various Settings of Training Data. Paper presented at the *The Ninth NTCIR Workshop on Evaluation of Information Access Technologies: Information Retrieval, Question Answering, and Cross-Lingual Information Access*, Tokyo, Japan.
- Tseng, Y. H., & Tsay, M. Y. (2013). Journal clustering of Library and Information Science for subfield delineation using the bibliometric analysis toolkit: CATAR. *Scientometrics*, 95(2), 503-528.
- Urban-Lurain, M., Prevost, L., Haudek, K. C., Henry, E. N., Berry, M., & Merrill, J. E. (2013). Using computerized lexical analysis of student writing to support Just-in-Time teaching in large enrollment STEM courses. Paper presented at the *Frontiers in Education Conference, 2013 IEEE*, Oklahoma City, OK.
- Vo, D. T., & Ock, C. Y. (2015). Learning to classify short text from scientific documents using topic models with various types of knowledge. *Expert Systems With Applications*, 42(3), 1684-1698. doi:10.1016/j.eswa.2014.09.031
- Wallach, H. M. (2006). Topic modeling: beyond bag-of-words. Paper presented at the *Proceedings of the 23rd international conference on Machine learning*, Pittsburgh, PA. Retrived from <http://dl.acm.org/citation.cfm?doid=1143844.1143967>
- Yang, Z., Fan, K. F., Lai, Y. X., Gao, K. M., & Wang, Y. (2014). Short Texts Classification Through Reference Document Expansion. *Chinese Journal of Electronics*, 23(2), 315-321.
- Zehner, F., Sälzer, C., & Goldhammer, F. (2015). Automatic Coding of Short Text Responses via Clustering in Educational Assessment.

*Educational and Psychological Measurement*, 76(2), 280-303.

doi:10.1177/0013164415590022

Zuo, Y., Zhao, J., & Xu, K. (2015). Word network topic model: a simple but general solution for short and imbalanced texts. *Knowledge and Information Systems*, 48, 379-398. doi:10.1007/s10115-015-0882-z

## A Preliminary Study of Short Text Responses Clustering for Mobile Educational Activities

**Lung-Hao Lee**

Postdoctoral Fellow  
Graduate Institute of Library and Information Studies  
National Taiwan Normal University

**Yu-Ta Chien**

Postdoctoral Fellow  
Science Education Center  
National Taiwan Normal University

**Chun-Yen Chang**

Chair Professor  
Graduate Institute of Science Education  
National Taiwan Normal University

**Tsung-Yen Li**

Research Assistant  
Science Education Center  
National Taiwan Normal University

**Yuen-Hsien Tseng**

Professor  
Graduate Institute of Library and Information Studies  
National Taiwan Normal University

### Introduction

Text clustering is a traditional information retrieval technique to

detect topics and their relations for information browsing, analysis, and organization in Library and Information Science. On the one hand, the techniques of Information Retrieval (IR) and/or Natural Language Processing (NLP) have advanced to a level that affects the research landscape of academic domains and has practical applications in various industrial sectors. On the other hand new learning tools or teaching paradigms have also changed in-class interactions, such as with the use of Instant Response Systems (IRS). The advance of these two fields have converged to support some online or on-site course activities that were previously infeasible, like real-time understanding of student's responses.

IRS has been tested and used in higher education classrooms since the 1960's. Recently, IRS has gained greater popularity in class interaction. Throughout 2014, domestic IRS services in Taiwan like Zuvio usage at the National Taiwan University (NTU) increased from 61 to 263 instructors, 68 to 384 courses, and 2,037 to 11,172 students participating.

By broadcasting a question to all students' mobile devices and obtaining responses instantly, such systems help teachers elucidate the learning status of each student and also help students maintain their attention during class as a result of the instant feedback from teachers and classmates. However, the potential of such IRS may still be under-explored. In the NTU case, using traditional Chinese text, the majority (54%) of questions deployed in Zuvio were multiple choices, though instructors also used open-ended questions (20%) and composite questions (21%) to promote deeper engagement and reflection. Previous studies indicated that multiple-choice examinations pose an obstacle for higher-level thinking in science classes and constructed response (e.g. free text writing) assessments are widely viewed as providing greater insight into student thinking than closed form (e.g. multiple-choice) assessments. However, to our knowledge, no IRS system has yet provided analysis of these open-ended text responses in real time.

Based on the above trends and observations, this paper focuses on short text response processing in the situation where forms of IRS are used in and after class.

The above IR/NLP techniques can be applied, at least, to the following pedagogical and learning scenarios: in class, the teacher broadcasts a question to the students' mobile devices and obtains nearly immediate short

text responses. The IR/NLP techniques cluster the short texts in real time for the teacher to browse through the organized texts in an effective manner to better know the students' responses and concomitant learning status. Or, the students are grouped together for peer discussion/debate about the question based on their clustered text responses. This kind of classroom interaction has been shown to be an effective pedagogy using Yes/No questions and is expected to be more effective if students' responses are open to any thoughts, and not limited to the teacher's pre-determined choices. Also, lexical analysis can be performed for each cluster to generate a diagnosis report (showing salient terms and term co-occurrence networks in each cluster) to support "just-in-time teaching".

## Current Development

The short text responses from IRS pose new challenges to traditional text processing, namely, 1) insufficient statistical information, 2) new terms or abbreviations, 3) synonymy, 4) polysemy, 5) misspelling, and 6) atypical grammar.

Two strategies can be summarized from the literature to deal with the above challenges: 1) enriching the short texts with extra/external resources, and 2) seeking better models for semantic processing of short texts.

For the first strategy, Banerjee, Ramanathan, and Gupta (2007) proposed to cluster Google news by incorporating titles of the top-relevant Wikipedia articles as extra term features. Hu, Sun, Zhang, and Chua (2009) clustered short texts (Google snippets) by first extracting important phrases and then expanding the feature space by adding semantically close terms from WordNet and Wikipedia.

For the second strategy, Vector Space Model (VSM) is a classic approach to text processing. The semantic-bearing terms, also called features, from all the documents (short texts or ordinary articles) constitute the feature dimensions of the document vector space. Each document corresponds to a vector and the value of each dimension of the vector is often the term frequency (TF) multiplied by the inverse document frequency (IDF) of the term, e.g.,  $TF*IDF$ , or a normalization version of  $TF*IDF$ . Although VSM facilitates similarity computation between documents, it suffers from the vocabulary mismatch problem where exact terms occurring

in two documents are a requirement to build up the similarity between the two documents. Later development led to Latent Semantic Analysis (LSA) which uses singular value decomposition (SVD) of the term-by-document matrix to transform the original sparse feature space into a condense one that captures the topics of the terms and the documents simultaneously, and therefore alleviates the vocabulary mismatch problem.

Further development led to a category of probabilistic models which improved on the previous models. Probabilistic latent semantic indexing (pLSI) allows each document to exhibit multiple topics. The drawback of pLSI is that it allows no natural way to assign probabilities to previously unseen documents, making incremental indexing (processing of new documents) difficult. To address this issue, latent Dirichlet allocation (LDA) was developed. LDA models each document as a mixture of various topics, which is similar to pLSI, except that the topic distribution in LDA is assumed to have a Dirichlet prior.

## Real Case Application

To have an idea of how well existing clustering techniques can do for these short texts, we have tried three approaches: 1) Hierarchical Agglomerative Clustering (HAC) based on a hybrid way of term indexing, namely lexicon-based segmentation followed by a keyword extraction using the method of Tseng (2002), implemented in a well-debugged tool called CATAR; 2) Latent Semantic Analysis (LSA) based on the word segmentation by jieba and a topic modeling tool gensim without removal of any stopwords and punctuations; and 3) Latent Dirichlet Analysis (LDA) by jieba and gensim with stopwords and punctuations removed.

The result based on HAC is generally reasonable, and only a few texts could not be grouped together with other similar texts. This is because a rigorous criterion is imposed on the HAC, i.e., complete linkage clustering.

To improve the performance such that the texts containing semantically related terms are clustered together, it seems that LSA or LDA are better solutions. However, the results show that LSA and LDA cannot alone solve these problems better. They sometimes even lead to worse results. In addition to the shortage of textual information (short texts), there are other factors that influence the performance, such as textual style (whether

to remove stopwords and punctuation, or not), feature extraction (whether to use 1-grams as features in Chinese short texts or not), term expansion (whether to incorporate the term-level similarity into text clustering). Furthermore, these decisions may depend on the characteristics of the questions asked or classes taught.

## Conclusions

Text clustering is a powerful information retrieval technique to detect topics from document corpora, so as to provide information browsing, analysis, and organization. On the other hand, the Instant Response System (IRS) has been widely used in recent years to enhance student engagement in class and thus improve their learning efficacy and effectiveness. However, the lack of functions to process short text responses from the IRS prevents the further application of IRS in classes. Therefore, this article aims to discuss the characteristics of short text responses from the IRS, review current short text processing techniques, and demonstrate our implemented techniques through real-world examples in Taiwan, so as to provide experiences and insights for further study. In particular, we have compared three clustering methods for short traditional Chinese texts, and the result shows that theoretically better methods need not lead to better results, as there are various factors that may affect the final performance.

## Reference

- 曾元顯 (2002)。文件主題自動分類成效因素探討。中國圖書館學會會報, 68, 62-83。【Tseng, Y. H. (2002). Effectiveness Issues in Automatic Text Categorization. *Bulletin of the Library Association of China*, 68, 62-83.】
- 曾元顯、王峻禧 (2007)。分類不一致之自動偵測—以農資中心資料為例。圖書館學與資訊科學, 33 (2), 19-32。【Tseng, Y. H., & Wang, C. S. (2007). *Automatic Inconsistency Detection for the ASIC Categorization Collection*, 33(2), 19-32.】
- Banerjee, S., Ramanathan, K., & Gupta, A. (2007). *Clustering short texts using wikipedia*. Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, 787-788. doi: 10.1145/1277741.1277909

- Bartsch, R. A., & Murphy, W. (2011). Examining the Effects of an Electronic Classroom Response System on Student Engagement and Performance. *Journal of Educational Computing Research*, 44(1), 25-33. doi:10.2190/EC.44.1.b
- Beatty, I., & Gerace, W. (2009). Technology-Enhanced Formative Assessment: A Research-Based Pedagogy for Teaching Science with Classroom Response Technology. *Journal of Science Education and Technology*, 18(2), 146-162. doi:10.1007/s10956-008-9140-4
- Beatty, I. D., Gerace, W. J., Leonard, W. J., & Dufresne, R. J. (2006). Designing effective questions for classroom response system teaching. *American Journal of Physics*, 74(1), 31-39. doi:http://dx.doi.org/10.1119/1.2121753
- Birenbaum, M., & Tatsuoka, K. K. (1987). Open-Ended Versus Multiple-Choice Response Formats—It Does Make a Difference for Diagnostic Purposes. *Applied Psychological Measurement*, 11(4), 385-395. doi:10.1177/014662168701100404
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Caldwell, J. E. (2007). Clickers in the large classroom: current research and best-practice tips. *CBE-Life Science Education*, 6(1), 9-20. doi:10.1187/cbe.06-12-0205
- Cardoso, W. (2010). Clickers in foreign language teaching: a case study. *Contact: Teachers of English as a Second Language of Ontario*, 36(2), 36-55.
- Chen, G. B. (2015). *Word Co-occurrence Augmented Topic Model in Short Text*. (Unpublished master's thesis), National Cheng Kung University, Tainan, Taiwan.
- Chen, T. L., Lin, Y. F., Liu, Y. L., Yueh, H. P., Sheen, H. J., & Lin, W. J. (2013). Integrating Instant Response System (IRS) as an In-Class Assessment Tool into Undergraduate Chemistry Learning Experience: Student Perceptions and Performance. In M. H. Chiu, H. L. Tuan, H. K. Wu, J. W. Lin, & C. C. Chou (Eds.), *Chemistry Education and Sustainability in the Global Age* (pp. 267-275). Netherlands: Springer.
- Cheng, X. Q., Yan, X. H., Lan, Y. Y., & Guo, J. F. (2014). BTM: Topic Modeling over Short Texts. *Ieee Transactions on Knowledge and Data*

- Engineering*, 26(12), 2928-2941. doi:10.1109/tkde.2014.2313872
- Chien, Y. T., & Chang, C. Y. (2015a). *Providing students with an alternative way to interact with the teacher in the silent classroom: Teaching with the CloudClassRoom technology*. Poster session presented at the Inaugural Asian Conference on Education & International Development (ACEID), Osaka, Japan.
- Chien, Y. T., & Chang, C. Y. (2015b). Supporting socio-scientific argumentation in the classroom through automatic group formation based on students' real-time responses. In M. S. Khine (Ed.), *Science education in East Asia: Pedagogical innovations and research-informed practices* (pp. 549-563). Springer: International Publishing.
- Christopherson, K. M. (2011). Hardware or wetware: what are the possible interactions of pedagogy and technology in the classroom? *Teaching of Psychology*, 38, 288-292. doi:10.1177/0098628311421332
- Chuang, Z. J., & Tseng, Y. H. (2008). NTCIR-7 Experiments in Patent Translation based on Open Source Statistical Machine Translation Tools. Paper presented at the *Seventh NTCIR Workshop on Evaluation of Information Access Technologies: Information Retrieval, Question Answering, and Cross-Lingual Information Access*, Tokyo, Japan.
- Deal, A. (2007). *A Teaching with Technology White Paper: Classroom Response Systems*. Retrieved from [https://www.cmu.edu/teaching/technology/whitepapers/ClassroomResponse\\_Nov07.pdf](https://www.cmu.edu/teaching/technology/whitepapers/ClassroomResponse_Nov07.pdf)
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391-407. doi:10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASII>3.0.CO;2-9
- Gilbert, A. (2005). *New for back-to-school: 'Clickers'*. Retrieved from <http://www.cnet.com/news/new-for-back-to-school-clickers/>
- Han, J. H. (2014). Closing the Missing Links and Opening the Relationships among the Factors: A Literature Review on the Use of Clicker Technology Using the 3P Model. *Journal of Educational Technology & Society*, 17(4), 150-168.
- Hearst, M. A. (2015). Can Natural Language Processing Become Natural Language Coaching? Paper presented at *The 53rd Annual Meeting of the Association of Computational Linguistics*, Beijing, China.

- Hoffman, J. (2012). *Speak Up? Raise Your Hand? That May No Longer Be Necessary*. Retrieved from [http://www.nytimes.com/2012/03/31/us/clickers-offer-instant-interactions-in-more-venues.html?\\_r=0](http://www.nytimes.com/2012/03/31/us/clickers-offer-instant-interactions-in-more-venues.html?_r=0)
- Hofmann, T. (1999). *Probabilistic latent semantic indexing*. Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, 50-57. doi: 10.1145/312624.312649 Retrieved from <http://dl.acm.org/citation.cfm?doid=312624.312649>
- Hu, X., Sun, N., Zhang, C., & Chua, T. S. (2009). *Exploiting internal and external semantics for the clustering of short texts using world knowledge*. Proceedings of the 18th ACM conference on Information and knowledge management, 919-928. doi: 10.1145/1645953.1646071 Retrieved from <http://dl.acm.org/citation.cfm?doid=1645953.1646071>
- Jameel, S., & Lam, W. (2013). *An unsupervised topic segmentation model incorporating word order*. Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, 203-212. doi: 10.1145/2484028.2484062
- Karakostas, A., Adam, D., Kioutsiouki, D., & Demetriadis, S. (2014). A pilot study of QuizIt: The new android classroom response system. Paper presented at the *2014 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL)*, Thessaloniki, Greece.
- Lee, J. W. S., & Shih, M. L. (2015). Teaching Practices for the Student Response System at National Taiwan University. *International Journal of Automation and Smart Technology*, 5(3), 145-150. doi:10.5875/ausmt.v5i3.862
- Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., ... Zhang, H. J. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, 34(1), 51-57. doi:10.1016/j.cedpsych.2008.04.002
- Morais, A., Barragués, J. I., & Guisasola, J. (2015). Using a Classroom Response System for promoting interaction to teaching Mathematics to large groups of undergraduate students. *Journal of Computers in Mathematics and Science Teaching*, 34(3), 249-271.

- Salton, G. (1989). *Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer*. Reading, MA: Addison-Wesley.
- Sanchez, C. G., Castro, F., Gomez, J. I., Tenllado, C., Chaver, D., & Lopez-Orozco, J. A. (2012). *OpenIRS-UCM: an open-source multi-platform for interactive response systems*. Proceedings of the 17th ACM annual conference on Innovation and technology in computer science education, 232-237. doi: 10.1145/2325296.2325352 Retrieved from <http://dl.acm.org/citation.cfm?doid=2325296.2325352>
- Siau, K., Hong, S., & Nah, F. F. H. (2006). Use of a classroom response system to enhance classroom interactivity. *IEEE Transactions on Education*, 49(3), 398-403. doi:10.1109/TE.2006.879802
- Stanger-Hall, K. F. (2012). Multiple-Choice Exams: An Obstacle for Higher-Level Thinking in Introductory Science Classes. *CBE Life Sciences Education*, 11(3), 294-306. doi:10.1187/cbe.11-11-0100
- Tang, J. L., Wang, X. F., Gao, H. J., Hu, X., & Liu, H. (2012). Enriching short text representation in microblog for clustering. *Frontiers of Computer Science*, 6(1), 88-101. doi:10.1007/s11704-011-1167-7
- Tseng, Y. H. (1998). *Multilingual Keyword Extraction for Term Suggestion*. Proceedings of the 21st International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '98, 377-378. doi: 10.1145/290941.291066
- Tseng, Y. H. (2002). Automatic Thesaurus Generation for Chinese Documents. *Journal of the American Society for Information Science and Technology*, 53(13), 1130-1138.
- Tseng, Y. H. (2010a). *Content Analysis Toolkit for Academic Research (CATAR)*. Retrieved from <http://web.ntnu.edu.tw/~samtseng/CATAR/>
- Tseng, Y. H. (2010b). Generic Title Labeling for Clustered Documents. *Expert Systems With Applications*, 37(3), 2247-2254.
- Tseng, Y. H., Chang, C. Y., Rundgren Chang, S. N., & Rundgren, C. J. (2010). Mining Concept Maps from News Stories for Measuring Civic Scientific Literacy in Media. *Computers & Education*, 55(1), 165-177. doi:10.1016/j.compedu.2010.01.002
- Tseng, Y. H., Liu, C. L., Tsai, C. C., Wang, J. P., Chuang, Y. H., & Jeng, J. (2011). Statistical Approaches to Patent Translation for PatentMT-

- Experiments with Various Settings of Training Data. Paper presented at the *The Ninth NTCIR Workshop on Evaluation of Information Access Technologies: Information Retrieval, Question Answering, and Cross-Lingual Information Access*, Tokyo, Japan.
- Tseng, Y. H., & Tsay, M. Y. (2013). Journal clustering of Library and Information Science for subfield delineation using the bibliometric analysis toolkit: CATAR. *Scientometrics*, 95(2), 503-528.
- Urban-Lurain, M., Prevost, L., Haudek, K. C., Henry, E. N., Berry, M., & Merrill, J. E. (2013). Using computerized lexical analysis of student writing to support Just-in-Time teaching in large enrollment STEM courses. Paper presented at the *Frontiers in Education Conference, 2013 IEEE*, Oklahoma City, OK.
- Vo, D. T., & Ock, C. Y. (2015). Learning to classify short text from scientific documents using topic models with various types of knowledge. *Expert Systems With Applications*, 42(3), 1684-1698. doi:10.1016/j.eswa.2014.09.031
- Wallach, H. M. (2006). Topic modeling: beyond bag-of-words. Paper presented at the *Proceedings of the 23rd international conference on Machine learning*, Pittsburgh, PA. Retrived from <http://dl.acm.org/citation.cfm?doid=1143844.1143967>
- Yang, Z., Fan, K. F., Lai, Y. X., Gao, K. M., & Wang, Y. (2014). Short Texts Classification Through Reference Document Expansion. *Chinese Journal of Electronics*, 23(2), 315-321.
- Zehner, F., Sälzer, C., & Goldhammer, F. (2015). Automatic Coding of Short Text Responses via Clustering in Educational Assessment. *Educational and Psychological Measurement*, 76(2), 280-303. doi:10.1177/0013164415590022
- Zuo, Y., Zhao, J., & Xu, K. (2015). Word network topic model: a simple but general solution for short and imbalanced texts. *Knowledge and Information Systems*, 48, 379-398. doi:10.1007/s10115-015-0882-z