

線上作文寫作訂正系統

鄭宇均 吳世弘 林承億
朝陽科技大學 朝陽科技大學 朝陽科技大學
g831147831147@gmail.com shwu@cyut.edu.tw mars60585@yahoo.com.tw

摘要

作文能力對於現在學生是不可或缺，作文有助學生的思想、溝通能力。我們相信加強每個學生的作文能力，有助於他未來的發展。考慮到現今社會網路的發達，人們對於網路的依賴性極高，於是我們決定朝網路這方面去幫助學生。本專題將作文寫作與網頁結合，做成一個寫作輔助網，可以藉由偵測錯別字、成語使用、作文格式、標點符號查詢等功能，來幫助學生加強作文能力。

關鍵詞：寫作輔助網

Abstract

Writing is an essential skill for students. Writing skill can benefit students on communication ability and thinking. We believe that helping students to improve the writing skill is investing their development in future. In the Internet era, people working are deeply depending on the network. We will support students writing on the network. Our system consists of a website for writing practice. As a writing support website, it can detect errors on word usage, also provide idiom inquire system and format checking system to help students.

Keywords: Writing Support Website

1 簡介

各個國家的語言教育，皆脫離不了聽、說、寫、讀這四個方面。而在這四個方面中，尤其以“寫”這一環節最為重要，作文是一個同時關乎「溝通能力」以及「邏輯能力」的訓練，能寫出一手好文章的人，溝通能力與邏輯能力必然不差。寫作不僅可以培養一個人的表達能力、文學素養，甚至可以激發、訓練一個人的組織與思考以及增進創造、理解等能力。

作文能力很重要，學生在作文訓練上需更加強。有鑑於傳統寫作的限制，造成的現今寫作的困境，我們希望利用網際網路的特性以突破以往的限制。只要有網路、電腦的地方就可

以馬上進行寫作學習，因此我們設計一套線上寫作系統，並附加輔助功能藉此精進學生的寫作能力。

寫作不是到學校才能練習的事情。透過網路與寫作的結合，可以培養學生語文領域方面的寫作能力。也可以在下課後或假日時，主動上網練習，精進自己的作文能力。所以決定發展網路化的寫作學習寫作環境，來輔助和加強學生的作文能力。

2 系統概述

2.1 系統介紹

本系統著重於網際網路化寫作取代傳統寫作，使得任何人都可以在有網路、電腦的情況下立馬練習寫作。現在由於網路的普及，導致人們對網路的依賴性極高。因此希望藉由此系統能使學生對寫作的練習慾望更加強，進而增加學生們的作文能力。

表一比較了傳統寫作跟網頁寫作的差異。此系統本身也有輔助寫作的功能。偵測系統方面有測錯別字、成語使用偵測、作文格式偵測，把這項功能包裝成 API，方便其他人使用。另外查詢功能方面有標點符號查詢、根據題目給予提示等。以上是寫作時的功能，寫完後也有文章收藏的功能。此系統的作文區查詢寫完的作文，並會把寫作過程中所偵測出的錯別字、成語使用、格式偵測和作文內容，皆以網頁來輸出，方便學生查閱。

	傳統作文	線上作文系統
文章保存性	低	高
錯別字、成語使用	透過字典查詢	即時偵測
寫作地方	紙張	網路上
寫作方法	紙、筆	透過網路
提示	只有題目說明	除了說明，還給予隨機提示

表 1：傳統作文與線上作文系統之比較

2.2 系統架構

以 PHP & HTML 設計出作文寫作網頁，其架構圖如下圖所示：

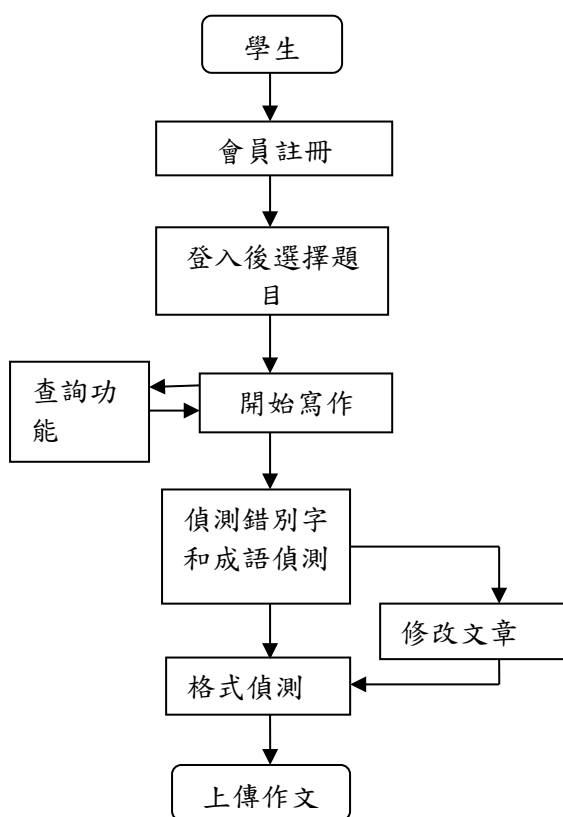


圖 1：系統流程圖

2.3 API 介紹

API 為應用程式介面 (Application Programming Interface，簡稱：API)，是指軟體系統不同組成部分銜接約定。

API 我們是應用在偵測功能系統，系統是安裝於朝陽科大 M-415 實驗室的電腦 Server 上。學生於作文網所輸入的作文，此系統的 Server 會將文章經過處理。將一串一串的文字包裝後透過 AJAX 來傳入到 M-415 的 Server 上，並取得回傳值。如圖 2 所示。

偵測功能是使用 AJAX 來達成目的，AJAX 是可以達到不更新頁面情況下而更新頁面上的資料或傳遞資料的語言。所有資料在傳遞過程中，都是由 JSON 來包裝做傳遞。

JSON 的格式重點是由物件(object)用大括號 { }，陣列(array)用中括號 [] 所組成的。因此取得回傳值後，必須將回傳值經過特殊處理，才能顯示出學生容易看的文字。

Server 間的傳遞方法是使用 POST，在網頁上市很常見。有另外一個方式是 GET，但 GET 會將帶有網頁表單的參數，導致網頁的資料容易外流，因此我們選擇 POST 來做傳遞。兩者間的差異如表 2 所示

	GET	POST
網址差異	會包含表單參數、資料。	傳遞時，網址不變。
資料傳遞量	有長度限制。	長度不受限。
安全性	表單參數、內容皆可不保密。	透過 HTTP Request 傳遞，保密性高。

表 2：GET 與 POST 之比較

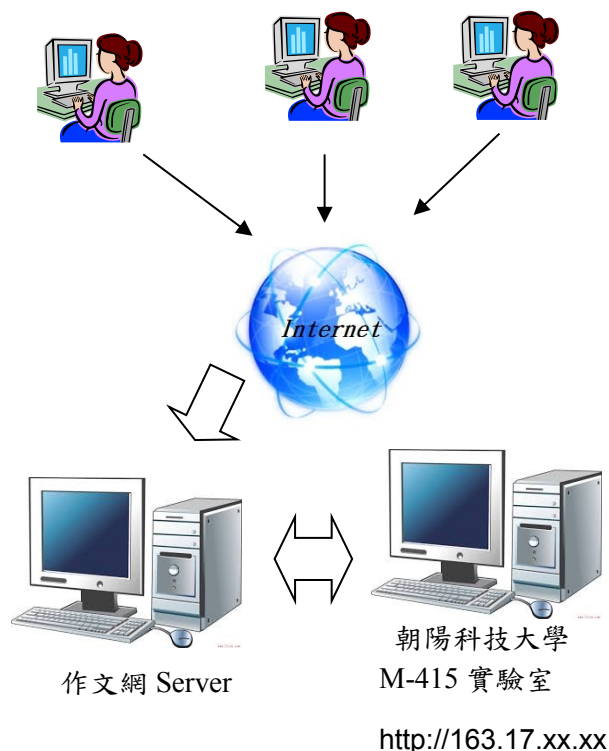


圖 2：偵測系統 API

AJAX 傳送程式：

```
$.ajax({ //AJAX 傳值到 API 程式連結 JAVA
  url: "api/API.php", //傳送目的地
  data: "&id="+s, //傳送資料
  dataType: 'json', //傳送資料的型態
  async: false, //同步 OR 異步
  type: "POST", //Server 間的傳送協定
  cache: false,
  success: function(data)
  { //傳送成功所執行的 code }}
```

API 的優點是可以跟主 Server 分開維護跟管理，也可以各自在進行新的開發。API 也能讓別人透過網址來使用，讓別人整合出更好的系統。

API SERVER 端接收程式：

```
<?php header('Content-  
type:application/json;charset=UTF-8');  
if (is_array($_GET)&&count($_GET)>0)  
{  
    if(isset($_GET["id"]))//是否存在"id"這個參數  
    {  
        $id=$_GET["id"];//存在  
        $Detection = "java Wrong ".urlencode($id);  
        $Output = exec($Detection);  
        $Output=urldecode($Output);  
        echo $Output ;  
    }  
} else if (is_array($_POST)&&count($_POST)>0) {  
    if(isset($_POST["id"]))//是否存在"id"這個參數  
    {  
        $id=$_POST["id"];//存在  
        $Detection = "java Wrong ".urlencode($id);  
        $Output = exec($Detection);  
        $Output=urldecode($Output);  
        echo json_encode($Output) ;  
    }  
} else  
{echo 'NO Input'; }  
?>
```

偵測系統本身是根據資料庫本身內的國中生作文來做比對搜尋，並判斷是否為錯別字。作文資料庫是利用 Java 裡的 Lucene 功能來建立的。Lucene 是可以建立一個屬於自己的索引引擎，在與資料庫的比對速度上來說是很快速的。Lucene 索引文件格式獨立於應用平台，使得兼容系統或者不同平台的應用能夠共享建立的索引文件。

2.4 系統功能

- (1) 錯別字偵測
- (2) 成語使用偵測
- (3) 作文格式偵測
- (4) 標點符號查詢
- (5) 根據題目給予提示

3. 成果展示

3.1 學生註冊

由於這是一個作文網，必須要有會員。學生透過帳號、密碼登入後，才能開始寫作、觀看歷史作文。管理者能清楚了解所有

會員的登入情形、寫作狀況。方便管理者做網頁上的管理。如圖 3 所示。

會員註冊畫面，包含帳號、密碼、學校、信箱、年級、性別等欄位，以及送出、清除、回到首頁按鈕。

圖 3：會員註冊畫

帳號、密碼、學校、信箱等欄位也有做格式判斷，當學生輸入的數值不符合規定時，也會告訴學生格式錯誤。如圖 4 所示。帳號有即時比對系統，避免有重複的帳號出現。如圖 5 所示。

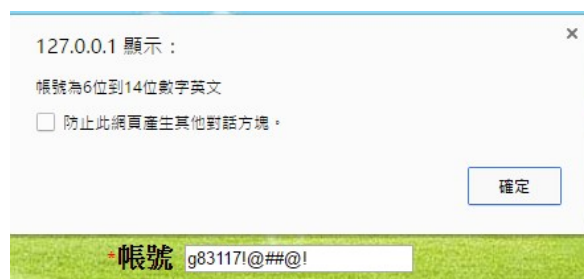
127.0.0.1 顯示：帳號為6位到14位數字英文。防止此網頁產生其他對話方塊。確定。下方顯示帳號欄位為 g83117!@##@!

圖 4：帳號格式錯誤警告

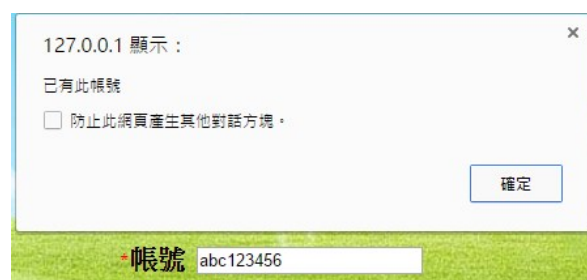
127.0.0.1 顯示：已有此帳號。防止此網頁產生其他對話方塊。確定。下方顯示帳號欄位為 abc123456

圖 5：帳號重複警告圖

3.2 題目選擇

此系統的作文題目，皆是國中會考、國中基測歷屆試題的題目。當選擇題目時，會出現相對應的題目解說。如圖 6 所示。

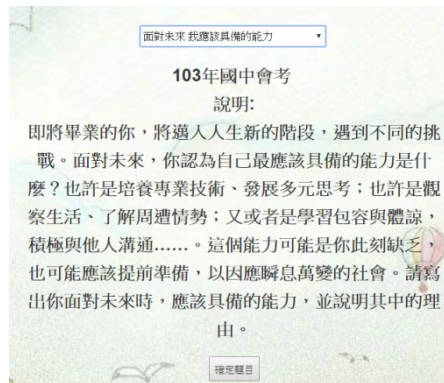


圖 6：題目說明

3.3 開始寫作

3.3.1 提示功能

當你開始寫作時，卻不知道如何下手時，可以點選提示功能。此功能會根據題目挑出幾個詞來做造句，讓使用者看完幾個句子，對於寫作有幫助。如圖 7 所示。

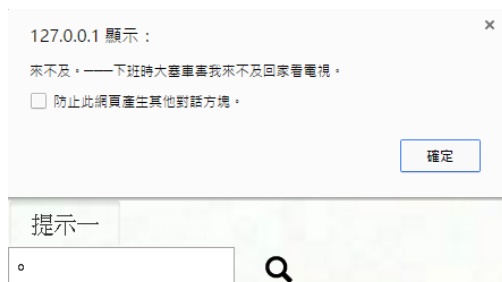


圖 7：提示功能

3.3.2 標點符號查詢

許多學生對於標點符號的使用不了解，每個符號都有它所代表的含意。如果胡亂使用的話，會導致一篇好的文章相形失色。所以我們有設計一個查詢標點符號的功能。如圖 8 所示。

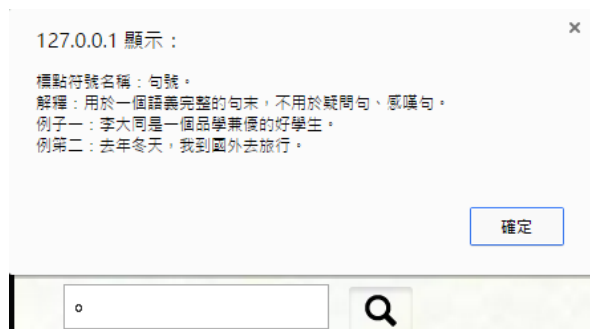


圖 8：標點符號查詢頁面的查詢。由

於標點符號有分半形跟全形，所以在查詢前必須先將所有半形轉換為全形。

3.3.3 錯別字偵測

將已寫好的文章，根據符號來分割成一串字串。並透過 AJAX 將資料傳送到 API Server 端，並取得回傳值列印出來。正確和錯誤會分別用藍色和紅色來區分。如圖 9 所示。

文字段落	錯別字	更正字	說明
由其這次的經驗對你來說，是很棒的經驗，	由其	尤其	
雖然說成績不怎麼好，下次比賽相信你一定能让對手甘拜下風，	成績	成績	
更重要的是這對你位來是一個很大的幫助啊！	位來	未來	
拜托了，	拜托	拜託	

圖 9：錯別字偵測圖

3.3.4 成語使用偵測

與錯別字一樣將文章分割後，傳送到 API Servers 端，並取得回傳值。正確和錯誤分別用藍色和紅色區分，並附有成語的釋義。如圖 10 所示。

文字段落	錯誤	正確	說明
但希望你能持之以恆繼續努力下去，		持之以恆	似始如一、有始有終半途而廢、虎頭蛇尾、一曝十寒有恆心的堅持到底。清·曾國藩 論紀澤：「爾之短處，在言語欠純訥，學止欠端重，看書不能深入，而作文不能崢嶸。若能從此三事上下一番苦工，進之以猛，持之以恆，不過一二年，自爾精進而不覺。」
下次比賽相信你一定能让對手甘拜下風，		甘拜下風	似心悅誠服、五體投地反不甘示弱、不甘雌服表示心悅誠服。鏡花緣第五十二回：「真是家學淵源，妹子甘拜下風。」三俠五義第三十九回：「他若真個為此而來，劣兄甘拜下風，從此後不稱御臨，也未為不可。」
讓自己的努力得以發揚光大，		發揚光大	良好的傳統、作風、事業在原有基礎上更加發展充實。如：「我國有悠久的歷史文化傳統，我們應該加以發揚光大。」
三人成虎。	三人成虎	三人成虎	戰國時，龐恭和魏太子同去趙國當人質，臨行前龐恭以三人成虎勸諫惠王不要聽信讒言。典出戰國策 魏策二。比喻謠言惑眾。隋唐演義 第二回：「正是積毀成山，三人成虎。到開皇二十年十月，隋主御武德殿，宣詔廢典為庶人。」亦作「市虎」。

圖 10：成語使用偵

3.3.5 文章格式偵測

當文章全部寫好時，我們還有格式偵測的功能。可以針對文章有沒有「一逗到底」的情況，給予學生建議。某些特定詞後面也有相對應的標點符號，例如：「吧」、「啊」後面應該要接驚嘆號，或是「嗎」、「呢」後面應該要接問號。如圖 11 所示。

錯誤句：念大家吧？
更正建議：\吧\後面應該使用驚嘆號。
錯誤句：的時光啊。
更正建議：\啊\後面應該使用驚嘆號。
一逗到底：不錯喔!沒有發現一逗到底的情形。

圖 11：文章格式偵測

3.4 文章讀閱

當學生們寫好作文後，都直接把作文內容、功能所偵測出的結果都存入資料庫。在作文區，有此會員的所有作文列表，包含寫作時間、寫作題目、寫作字數都有詳細紀錄。如圖 12 所示。

作文題目	字數	時間	詳細
當我和別人意見不同的時候	306	2016-11-30 00:43:53	展開文章
從陌生到熟悉	326	2016-11-30 00:43:37	展開文章
那一次，我自己做決定	123	2016-11-30 00:43:20	展開文章
來不及	174	2016-11-27 19:10:05	展開文章
來不及	174	2016-11-27 18:58:28	展開文章

圖 12：文章列表

可以選擇你想看的文章，點選展開文章，就會進入詳細的文章網頁。有包含作文原文、所有錯別字偵測出的錯誤、成語使用偵測和作文格式偵測。都經過處理，清楚明瞭的顯示於網頁上並作好區隔。如圖 13~16 所示。

文章編號：324

我知道這些事情對你來說有一定的難度，但希望你能持之以恆繼續努力下去，這樣你這些努力是值得的，由其這次的經驗對你來說是很棒的經驗，雖然說成績不怎麼好，但至少你有努力過。下次比賽相信你一定能讓對手甘拜下風，拿下冠軍。讓自己的努力得以發揚光大，讓自己在高中生涯畫一個精彩的句點。老師也會為你感到驕傲的，你父母也會以你為榮。更重要的是這對你來說是一個很大的幫助啊。拜托了，同學，三人城虎。

圖 13：作文原文

錯別字偵測

句子	錯誤字	正確字
由其這次的經驗對你來說是很棒的經驗，	由其	尤其
雖然說成績不怎麼好，	成績	成績
更重要的是這對你來說是一個很大的幫助啊。	位來	未來
拜托了，	拜托	拜託
下次比賽相信你一定能讓對手甘拜下風，	甘敗下風	甘拜下風

圖 14：錯別字偵測顯

成語使用

句子	正確成語	錯誤成語	釋義
但希望你能努力不懈、持之以恆，	持之以恆		創始於和一一、有始有終，半途而廢，虎頭蛇尾，一曝十寒，有恆心的堅持到底。清·曹雪芹《紅樓夢》：「罷之短處，在言語欠條理，學識欠純正，事事不能深入，而作文不能條理，若能從此三事上下一番苦工，進之以恆，持之以恆，不過一二年，自可轉弱而為強。」
讓自己的努力得以發揚光大，	發揚光大		凡有好的傳統、作風、事業在原有基礎上更加發展充實。如：「我國有悠久的歷史文化傳統，我們應該加以發揚光大。」
下次比賽相信你一定能讓對手甘拜下風，	甘拜下風		真心佩服。五體投地，甘拜下風。不甘願服輸。中心佩服。陳乃鋒 第三十二回：「真是甘拜下風，勝了甘拜下風。」三傳五典 第三十九回：「他若真個為此事而來，必是甘拜下風，從此決不稱強，也未為不可。」
三人城虎，	三人成虎	三人成虎	第三回：「真帝和魏太子同去趙國家人宴，臨行前魏帝以三人成虎勸魏王不要輕信讒言。典出戰國策 魏策二。比喻謠言愈眾，聽者愈信。第二回：「正是韓昭伯山，三人成虎。劉備二十年十月，魏主派使臣，宣詔魏王為人。」亦作「市虎」。

圖 15：成語使用偵測顯

示

格式偵測

錯誤句：念大家嗎！更正建議：\嗎\後面建議使用問號
錯誤句：的時光啊！更正建議：\啊\後面應該使用驚嘆號
一逗到底：不錯喔!沒有發現一逗到底的情形

圖 16：格式使用偵測顯

4.分析與檢討

我們針對錯別字偵測上做了正確率的測試，我們收集了國中生寫作並請老師批改過的作文。抽出 20 篇作文來做測試。總共有 38 個人工改錯的錯誤，偵測系統偵測出 24 個錯誤，其中有 10 個是跟人工錯誤相符合的。另外 12 個錯誤是偵測不正確，還有偵測出 2 個是人工所漏掉的錯誤。系統漏偵測的錯誤有 28 個。如表 3 所示。

表 3：錯別字偵測測試

	人工改錯	系統偵錯
總錯誤	38	24
偵測不正確		12
偵測正確		10
系統漏偵		28
人工遺漏	2	

由表 3 發現偵測系統完整度還不夠好，未來希望能增加資料庫的作文量，使得偵測系統的偵測率跟正確率能上升。判斷錯別字時，判斷演算法能再加強，使得誤判錯別字率降低。根據某些詞性來判斷錯別字，例如：我「做」得，做本身詞性是動詞，因此後面必須接「得」而不是「的」。

在系統上出現蠻多誤判的情形，例如：然後我「就想」到要讓人快樂，在系統上會把「就想」判斷為「就像」，導致判斷錯誤。所以偵測程式對於錯別字的判斷應該再加強，才不會出現更多誤判的情況。

5. 結論與未來展望

5.1 結論

有鑑於現在年輕學生都使用手機、電腦、平板等，打字多於手寫。此系統可以解決現在學生懶得寫字的問題，而且還能幫助學生更正錯字，可以邊寫作邊了解自己的錯誤。希望藉由此系統能幫助學生克服對寫作文的恐懼，也能增加學生寫作經驗。讓學生在面對各種題目時能文思泉湧不在躊躇不前，更藉由此練習系統減少在寫作文的時候會常發生的錯誤，讓學生能更輕鬆地取得高分，更希望能藉此增加我國學生作文能力。

5.2 未來展望

目前本系統著重於輔助學生寫作，以及增進學生寫作時的方便性。但我們認為要全面性提升寫作能力，只憑這些是不夠的。

此系統的偵測率還有待加強，我們會再抓取更多作文來加強資料庫，才能讓成語、錯別字偵測率增高，錯誤率減低。另外希望在判斷錯別字上也能處理得更好，盡量減少出現判斷錯誤的地方。我們希望能夠增加更多輔助功能，如搭配詞選擇功能(提供更多同義字選擇或是其他詞彙組合)。以減少文章的錯誤用詞，並降低冗言贅字的頻率。

更期許未來能加上自動改錯，將偵測出的錯別字，直接於寫作欄位上更新，更人性化、便利的功能。

參考文獻

- [1]. 黃郁婷 著，網路互動寫作輔助系統，嘉義大學碩士論文，民國九十二年
http://www.tahrd.ntnu.edu.tw/files/archive/161_69f995ee.pdf
- [2]. 余思翰 著，中文作文寫作輔助系統，國立交通大學碩士論文，民國九十七年
<http://ndltd.ncl.edu.tw/cgi-bin/gs32/gsweb.cgi/login?o=dnclcdr&s=id=%22096NCTU5394027%22.&searchmode=basic3>. JavaScript
- [3]. Java 程式語言查詢

<http://www.w3schools.com/js/>

[4]. PHP 程式語言查詢

<http://www.w3schools.com/php/>

[5]. AJAX 程式語言查詢

http://www.w3schools.com/xml/ajax_intro.asp

[6]. CSS&HTML 查詢

<http://www.thewebpractice.com/w3c/Style/Examples/011/firstcss-tr.html>

[7]. 基礎 CSS 美化

<https://kkbruce.tw/bs2/BaseCSS>

[8]. 國中教育會考題庫

<http://cap.ntnu.edu.tw/examination.html>

[9]. 丁柏元 著，線上智慧型文章生成輔助系統，國立交通大學碩士論文，民國九十八年
<http://ndltd.ncl.edu.tw/cgi-bin/gs32/gsweb.cgi/login?o=dnclcdr&s=id=%22097NCTU5641024%22.&searchmode=basic>

[10]. 吳世弘、許力仁、林華煒、廖培凱，辨識自然語言指令之文字蘊涵識別系統，朝陽科技大學資訊工程學系 2014 年 12 月

[11]. Hen-Hsen Huang, Jun-Jie Wang, and Hsin-Hsi Chen (accepted). “Implicit Opinion Analysis: Extraction and Polarity Labelling.” *Journal of the Association for Information Science and Technology*. 2016 年
<https://www.aclweb.org/anthology/P/P16/P16-2004.pdf>

[12]. Han-Bin Chen, Hen-Hsen Huang, An-Chang Hsieh, Hsin-Hsi Chen (accepted). “A Simplification-Translation-Restoration Framework for Domain Adaptation in Statistical Machine Translation: A Case Study in Medical Record Translation.” *Computer Speech and Language*. 2016 年

<https://pdfs.semanticscholar.org/7b93/3815baf01d3d1a6b02cd8eb2d8746415d47e.pdf>

[13]. Write to Learn 線上系統英文寫作系統
<http://www.writetolearn.net/>

[14]. 彭淑惠、林佩樺 著，華語寫作測驗線上評分系統之運用，2012 年 5 月
http://www.sc-top.org.tw/download/research/TOCFL%20Writing_01.pdf