

# App Inventor2 實作網頁解析 翻譯機

大學部人工智慧應用  
數位系三年級

---

吳智鴻

教學網站：[HTTP://120.108.221.55/PROFCHWU/DCTEC](http://120.108.221.55/PROFCHWU/DCTEC)

FB社團： 106 數位系人工智慧

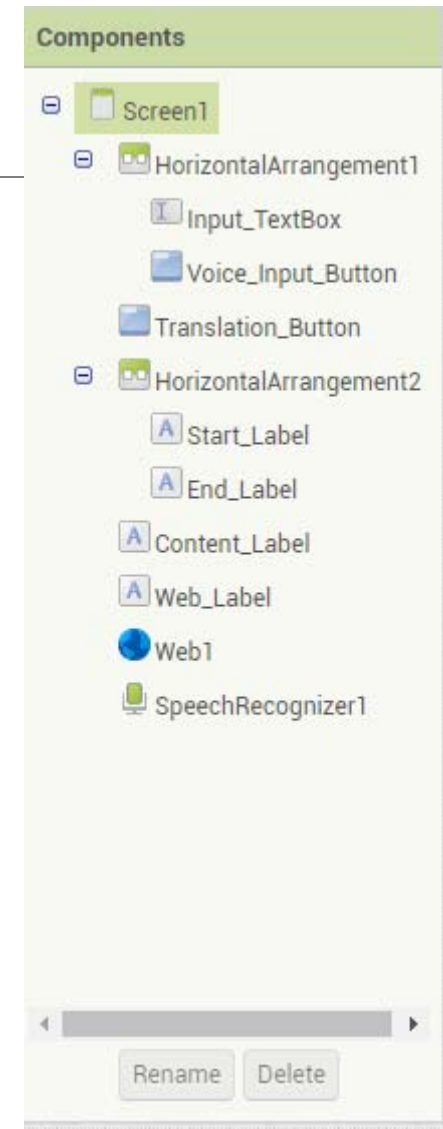
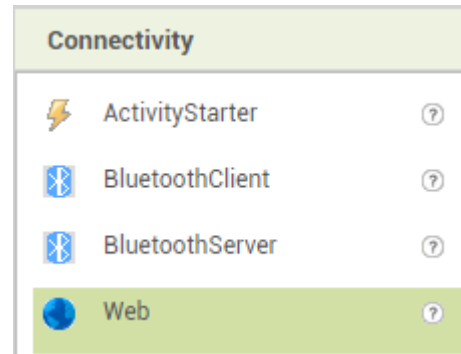
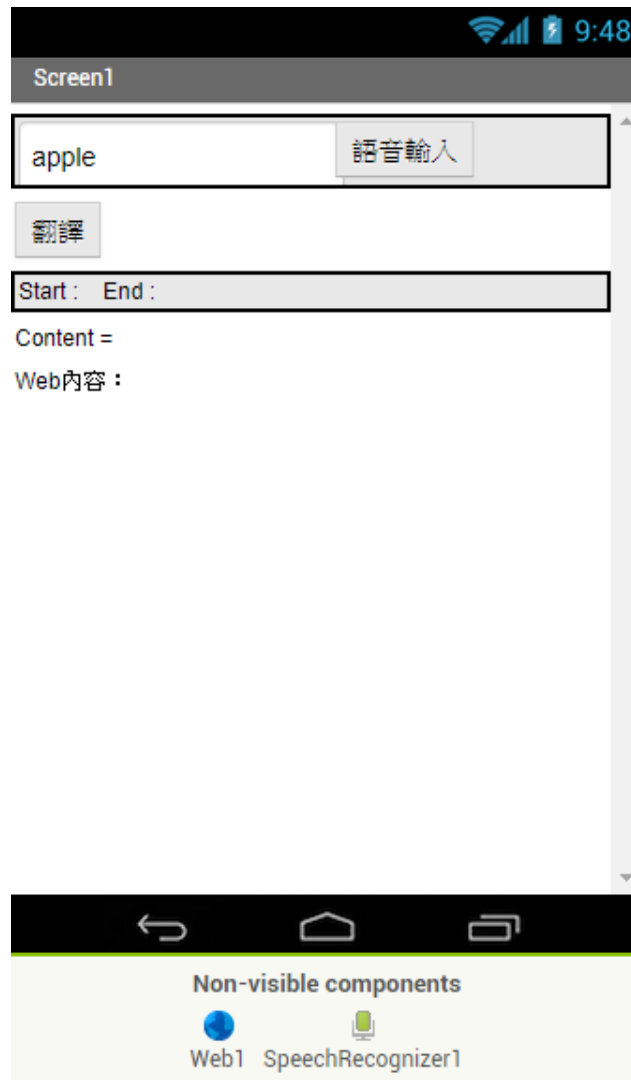
APP INVENTOR PROJECT: AITOLS (TRANSLATION)



# 螢幕設計



# 螢幕介面設計



# 步驟

---

## 1. 語音辨識輸入部分的處理

```
when Voice_Input_Button .Click  
do call SpeechRecognizer1 .GetText
```

```
when SpeechRecognizer1 .AfterGettingText  
  result  
do set Input_TextBox .Text to get result
```

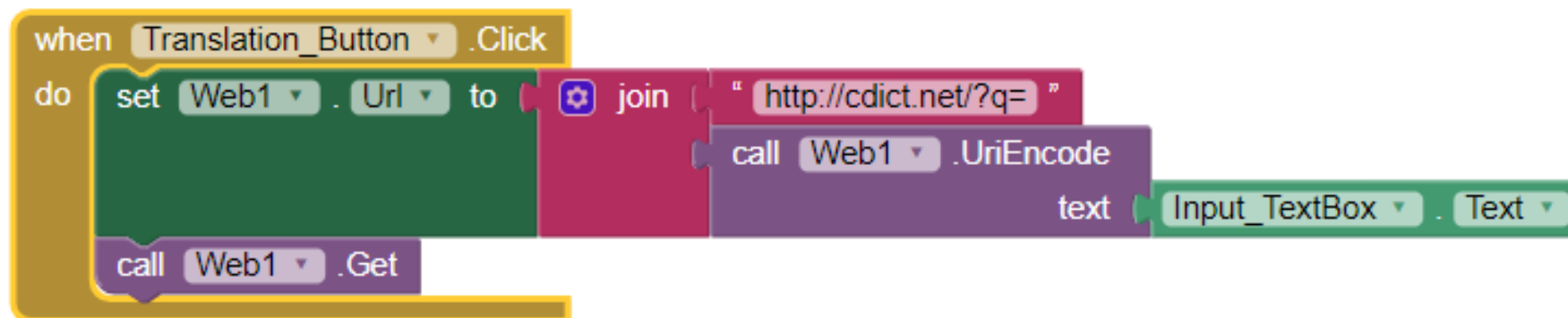
# 步驟

---

## 2. 用Web 元件抓取網頁

例如<http://cdict.net/?q=XXXX> XXX表示你要查詢的字串

記得搜尋的關鍵字要用UriEncode重新編碼



# 步驟3 解析網頁

---

解析網頁的意思，就是從一堆網頁HTML中，抓取出我所需要的部分。需要

1. 找到前置字元的所在位置（在網頁中的第幾個字元），例如：133
2. 從網頁中，從第133個字元，抓取 60個字元
3. 從這60個字元中，找到結尾字元的所在位置，例如: 55
4. 從這60個字元中，擷取所要的字元。

# 網頁擷取概念

```
initialize global start_location to 0
initialize global end_location to 0

when Web1 .GotText
  url responseCode responseType responseContent
do
  set Web_Label .Text to get responseContent
  set global start_location to starts at text get responseContent
  piece "content="
  set Start_Label .Text to join Start_Label .Text
  get global start_location
  if get global start_location ≠ 0
  then
    set Content_Label .Text to segment text get responseContent
    start get global start_location
    length 40
```

中華電信 4G 38% 21:33  
Screen3

翻譯

apple 語音輸入

翻譯

抓取的起始位置

start:123

content="apple 蘋果,蘋果電腦公司;人,傢夥;

原始網頁

```
<!DOCTYPE html>
<html lang="zh">
<head>
<meta charset="utf-8">
<title>apple 的中文翻譯 | 英漢字典</title>
<meta name="description"
content="apple 蘋果,蘋果電腦公司;人,傢夥;手榴彈">
<meta name="keywords"
content="apple,中文,翻譯,英漢,字典,中翻英,chinese,dictionary">
<meta property="og:title"
```

需要抓的地方



# 網頁擷取概念 # 2

開始字串

結束字串

initialize global start\_str to "content=" initialize global end\_str to ">"

```
when Web1 .GotText
  url responseCode responseType responseContent
do
  set Web_Label .Text to get responseContent
  set global start_location to starts at text get responseContent
  piece "content="
  set Start_Label .Text to join "start:"
  get global start_location
  if get global start_location ≠ 0
  then
    set Content_Label .Text to segment text get responseContent
    start get global start_location
    length 40
  set global end_location to starts at text Content_Label .Text
  piece get global end_str
  set End_Label .Text to join "end:"
  get global end_location
```

翻譯

apple 語音輸入

翻譯

start: 123 end: 26  
apple 蘋果,蘋果電腦公司;人,傢夥;手榴彈">  
<meta name="k  
<!DOCTYPE html>  
<html lang="zh">  
<head>  
<meta charset="utf-8">  
<title>apple 的中文翻譯 | 英漢字典</title>  
<meta name="description"  
content="apple 蘋果,蘋果電腦公司;人,傢夥;手榴彈">  
<meta name="keywords"  
content="apple,中文,翻譯,英漢,字典,中翻英,chinese,dictionary">





有了起始位置與結束位置，  
那應該怎樣把所需的翻譯結  
果取出呢？

---

想一想吧，自己試試看。



# 加入了擷取翻譯結果

```
when Web1 . GoText
  url responseCode responseType responseContent
do
  set Web_Label . Text to get responseContent
  set global_start_location to starts at text get responseContent
  piece "content="
  set Start_Label . Text to join "start:"
  get global_start_location
  if get global_start_location ≠ 0
  then
    set Content_Label . Text to segment text get responseContent
    start get global_start_location
    length 40
    set global_end_location to starts at text Content_Label . Text
    piece get global_end_str
    set End_Label . Text to join "end:"
    get global_end_location
    set Content_Label . Text to segment text Content_Label . Text
    start length get global_start_str + 2
    length get global_end_location - length get global_start_str + 2
```

```
Screen1
apple 語音輸入
翻譯
start : 123 end : 34
apple 蘋果,蘋果電腦公司;人,傢夥;手榴彈
<!DOCTYPE html>
<html lang="zh">
<head>
<meta charset="utf-8">
<title>apple 的中文翻譯 | 英漢字典</title>
<meta name="description" content="apple 蘋果,蘋果電腦公司;人,傢夥;手榴彈">
<meta name="keywords" content="apple,中文,翻譯,英漢,字典,中翻英,chinese,dictionary">
<meta property="og:title" content="apple 的中文翻譯 | 英漢字典">
<meta property="og:type" content="website">
<meta property="og:url"
content="http://cdict.net/q/apple">
<meta property="og:image" content="https://
assets.cdect.net/logo-100.png">
<meta property="og:site_name" content="英漢字典(CDICT)">
<meta property="fb:admins"
content="500282623">
<meta property="fb:app_id"
```



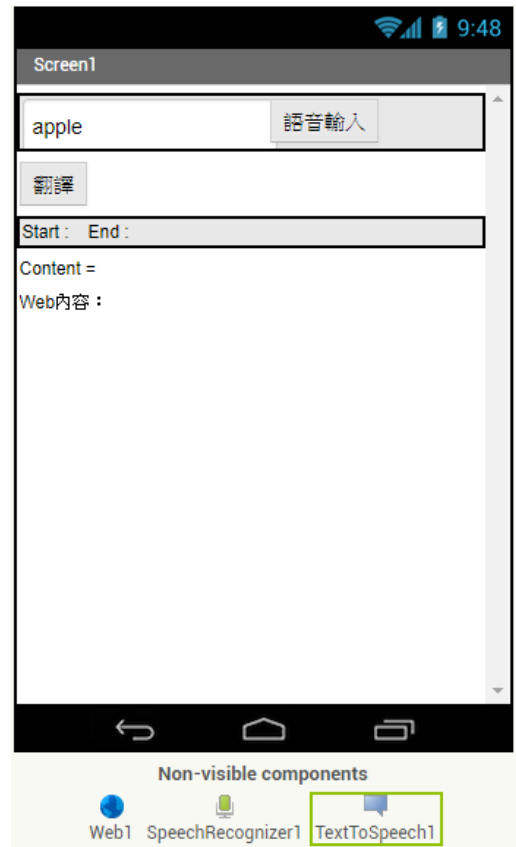
# 加上文字轉語音TTS功能吧

---

想一想吧，自己試試看。



# 加上文字轉語音功能



# 全部程式

```
initialize global start_location to 0
initialize global end_location to 0

when Web1 .GoText
  url responseCode responseType responseContent
do
  set Web_Label .Text to join ["Web="
                              "\n"
                              get responseContent ]
  set global start_location to starts at text get responseContent
  piece get global start_str
  set Start_Label .Text to join ["start:"
                                get global start_location ]
  if get global start_location ≠ 0
  then
    set Start_Label .Text to join ["start:"
                                  get global start_location ]
    set Content_Label .Text to segment text get responseContent
    start get global start_location + length get global start_str
    length 60
    set global end_location to starts at text Content_Label .Text
    piece get global end_str
    set End_Label .Text to join ["end:"
                                get global end_location ]
    set Content_Label .Text to segment text Content_Label .Text
    start 1
    length get global end_location - 1
    call TextToSpeech1 .Speak
    message Content_Label .Text
```

```
initialize global end_str to ">"
initialize global start_str to "content="

when Translation_Button .Click
do
  set Web1 .Uri to join ["https://cdict.net?q="
                        call Web1 .UriEncode
                        text Input_TextBox .Text ]
  call Web1 .Get

when voice_input_Button .Click
do
  call SpeechRecognizer1 .GetText

when SpeechRecognizer1 .AfterGettingText
result
do
  set Input_TextBox .Text to get result
```



# 加上第二套Yahoo字典

---

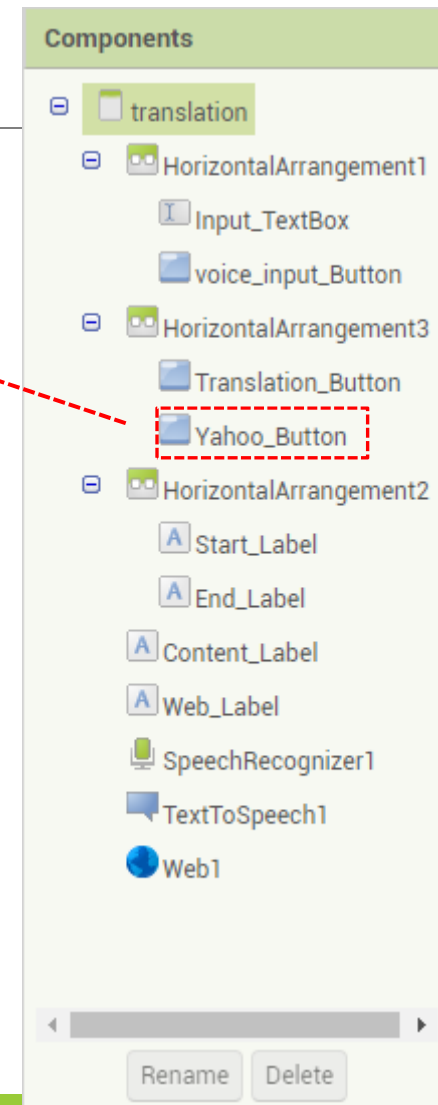
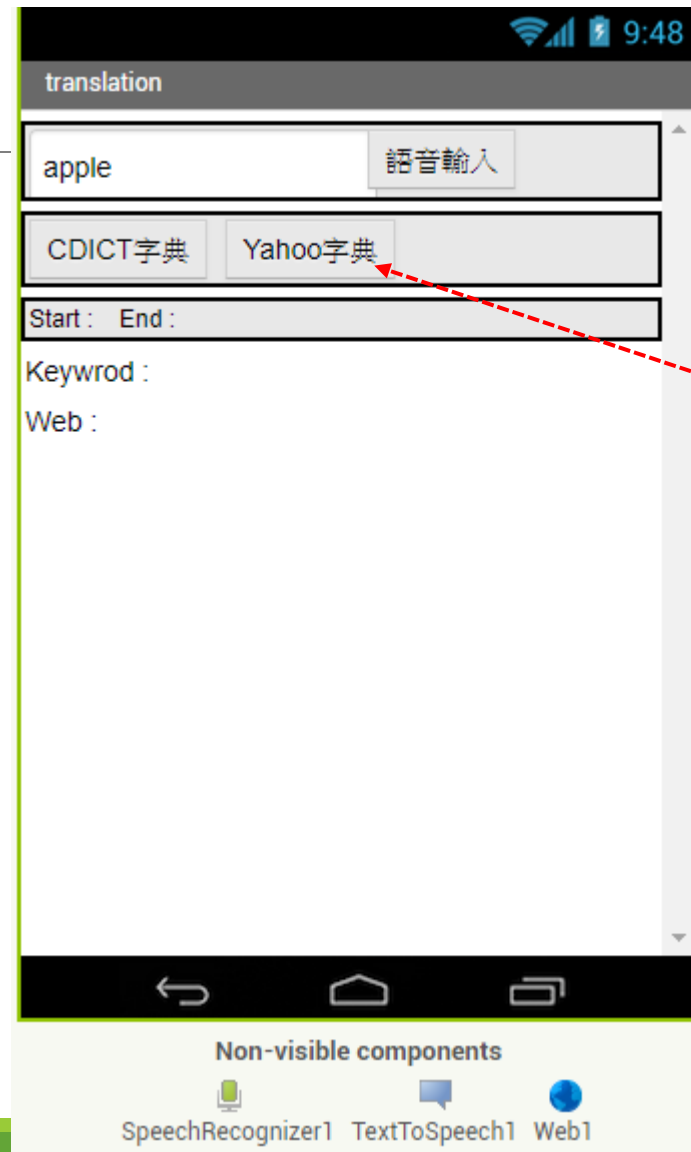
自行分析判斷YAHOO字典的網址?

他的開始字元?

它的結束字元?

# 螢幕設計

螢幕設計



# YAHOO字典的網頁擷取

定義YAHOO字典的開始與結束字元

```
when Yahoo_Button.Click
do
  set Web1.Url to join (" http://tw.dictionary.search.yahoo.com/search?p="
    call Web1.UriEncode
      text Input_TextBox.Text
  )
  set global dict_selection to 2
  set global start_str to " 1. "
  set global end_str to "< "
  call Web1.Get
```



# 執行畫面



start : 104411 end : 3

西瓜

Web=

```
<!DOCTYPE html>
<html id="Stencil" class="NoJs"
lang="zh-Hant-TW"><head><meta http-
equiv="content-type" content="text/
html; charset=UTF-8"><link rel="apple-
touch-icon" sizes="57x57" href="http:
//l.yimg.com/pv/static/img/yahoo
_purple_icon_iphone_60.png" /><link
rel="apple-touch-icon" sizes="114x114"
href="http://l.yimg.com/pv/static/img
/yahoo_purple_icon_iphone_120.png" /
><meta name="viewport"
content="width=device-width,maximum-
scale=1.0,initial-scale=1.0,minimum-
scale=0.9,user-scalable=no"><meta
http-equiv="x-dns-prefetch-control"
content="on"><noscript><meta HTTP-
```

# 加入YAHOO字典的全部程式

The image displays a collection of Scratch code blocks for a dictionary application. The blocks are organized into several functional sections:

- Global Initialization:** Four blocks at the top initialize global variables: `start_location` to 0, `end_location` to 0, `end_str` to ">", and `start_str` to "content=".
- Web1 GotText:** A large block that handles the initial web request. It sets `Web_Label` text, gets `responseContent`, and uses `starts at text` and `piece` blocks to find the start of the content. It then sets `Start_Label` text and uses an `if` block to check if `start_location` is not 0. If true, it sets `Start_Label` text, uses `segment text` to get a 60-character snippet starting from `start_location`, sets `global end_location` to the start of the next word, sets `End_Label` text, and uses `segment text` to get the next word starting from `1` up to `end_location`.
- voice\_input\_Button Click:** A block that calls `SpeechRecognizer1` when the voice input button is clicked.
- SpeechRecognizer1 AfterGettingText:** A block that sets the `Input_TextBox` text to the result from the speech recognizer.
- Translation\_Button Click:** A block that sets the `Web1` URI to a URL with the `Input_TextBox` text, sets `global dict_selection` to 1, sets `global start_str` to "content=" and `global end_str` to ">", and calls `Web1` .Get.
- Yahoo\_Button Click:** A block that sets the `Web1` URI to a Yahoo dictionary search URL with the `Input_TextBox` text, sets `global dict_selection` to 2, sets `global start_str` to "1." and `global end_str` to "<", and calls `Web1` .Get.