要怎麼讓電腦說話

Jing ([mqJing@gmail.com](mailto:mqJing@gmail.com))

要怎麼讓電腦說話? 你可以直接讓他播放一段預錄的語音 或者是讓電腦自動合成語音.今天要教大家的就是語音合成! 簡單的說, 就是輸入一篇文章給機器, 讓它能夠看著你的文章, 用人類的語言大聲地唸出來. 所以讀完本文, 任何人都可以讓你寫一個讓電腦說話的應用程式. 想不想讓你們家的6歲學齡前小孩, 聽聽看 Wikipedia 上綠蠵龜的解釋?

環顧一下四周, 你可以看到許多語音合成的應用. 例如: 汽車的導航系統, 電話語音自動回復, 盲人閱讀機. 語音合成也可以用在語音的教學上. 例如: 一段英文的念法. 互動遊戲的製作等. 值得一提的是, 世界上使用語音合成最有名的人物, 是英國著名物理學家 史蒂芬·霍金. 因為[肌肉萎縮性側索硬化症](http://zh.wikipedia.org/wiki/%E8%82%8C%E8%82%89%E8%90%8E%E7%B8%AE%E6%80%A7%E5%81%B4%E7%B4%A2%E7%A1%AC%E5%8C%96%E7%97%87)患者,全身癱瘓,不能發音.



[Stephen Hawking](http://en.wikipedia.org/wiki/Stephen_Hawking)

這個技術的主要優勢, 在於不用儲存聲音訊號, 取而代之的是使用合成語音軟體核心加上語音模型, 就可以用非常少的容量產生大量的語音資訊. 利用這項特性, 你就可以做出許多有創意的東西出來. 例如: 你可以製作一個玩具熊, 這玩具熊可以講 3000本以上童書的故事. 而且還可以模擬使用林志玲或馬英九總統的聲音, 來唸給你的孩子聽.

什麼叫做好的語音合成引擎? 研判的重點在於 **自然 (naturalness)** 與 **可理解 (intelligibility)** [1]. 比較知名引擎有 IBM tts (Text to Speech), Microsoft tts, Google tts, 我們國內的工研院也有研發類似的引擎[2-5]. 值得一提的是, 經過比較後, 工研院的中文發音聽起來異常親切而且免費.

今天我們來教大家怎麼使用 Microsoft 的服務, 簡單的說, 就是寫一個 application 利用微軟提供的語音合成服務, 讓它講中文. 但為什麼我們要學微軟的 solution 呢? 主要的想法是利用這次學習使用微軟的服務, 來學習將來微軟在 AZure Marketplace 電子市集的雲服務存取方式. 正是所謂的一石二鳥. 到時候, 真正想要用語音合成時, 再學工研院那介面的使用方式即可. 畢竟, 工研院的軟體服務與市集, 目前的規模還沒引起我的興趣.

這是一個 AP , 你會學到如何向微軟存取服務的 token. 利用這個 token 向微軟提出存取有關語音合成的功能函式呼叫權限, 達到語音合成的目的.

接下來, 開始我們的教學. 這份教學的大綱如下:

**Outline**

* Download (c++ 原始檔與完整專案下載)
* 如何註冊你的應用程式 on Azure Marketplace
* 開始寫程式

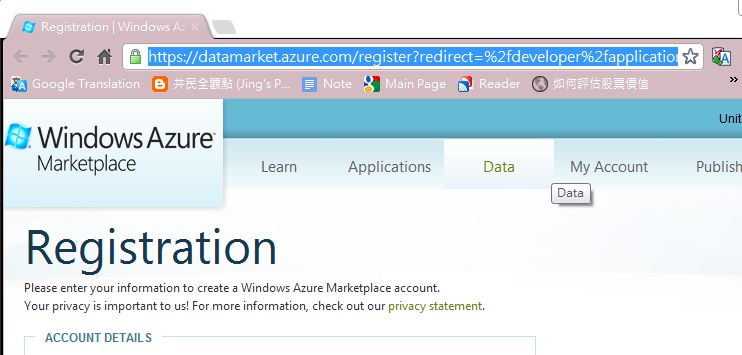
**Download**  (c++ 原始檔與完整專案下載)

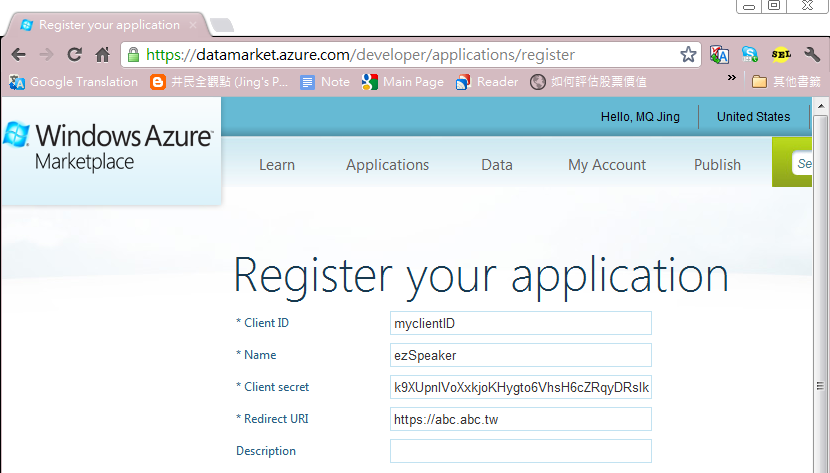
1. [cpp](https://docs.google.com/file/d/0B-3nWHvfq6BdNWdJRVpPZm1ZVUE/edit)
2. [project](https://docs.google.com/file/d/0B-3nWHvfq6BdVG00a3RKMmFOd2c/edit)

**如何註冊你的應用程式 on Azure Marketplace**

**Step** 1: 註冊 Azure Marketplace

* <https://datamarket.azure.com/developer/applications/>

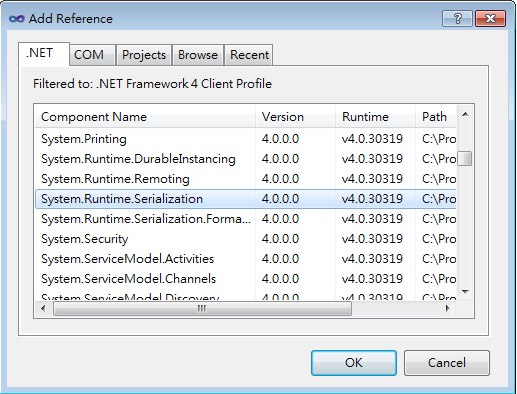




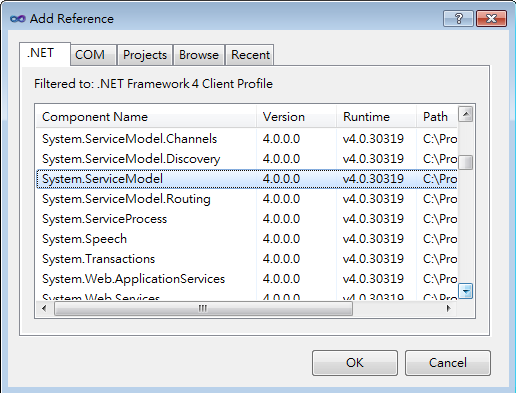
**開始寫程式**

**Step** 1: Add reference libraries. [Add Reference] -> [.NET]

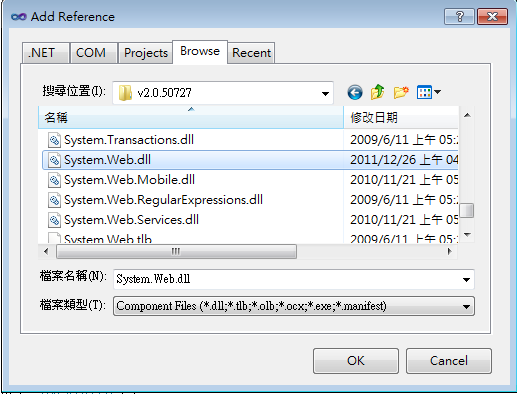
* System.Runtime.Serialization .dll for System.Runtime.Serialization.Json
* System.ServiceModel .dll for System.ServiceModel.Channels
* 把 [C:\Windows\Microsoft.NET\Framework\v2.0.50727\System.Web.dll] 加進來



System.Runtime Serialization



System.ServiceModel



System.Web (請使用 2.0.50727)

**Step** 2: 取得 authentication token

|  |
| --- |
| *// 查閱 client id & client secret:* [*https://datamarket.azure.com/developer/applications/*](https://datamarket.azure.com/developer/applications/)  String strClientID="<client id>"; *// 剛剛你在 azure marketplace 註冊的 client ID*  String strClientSecret = "<client secret>"; *// 你在 marketplace 上得到的對應密碼*  *// <關鍵片段> 取得 authentication token*  AdmAuthentication admAuth = new AdmAuthentication(strClientID, strClientSecret);  admToken = admAuth.GetAccessToken(); *// 取得 access token* |

// 輔助 class: 幫我們利用 Post 的方式, 向微軟取得 access token

|  |
| --- |
| // 申請窗口: <https://datamarket.accesscontrol.windows.net/v2/OAuth2-13>";  // 要求的權限 (scope): <http://api.microsofttranslator.com>  public class AdmAuthentication {  …  } |

**Step** 3: 使用 Speak method 搭配著 token, 要求說話.

|  |
| --- |
| *// Speak API 介面:* [*http://api.microsofttranslator.com/v2/Http.svc/Speak*](http://api.microsofttranslator.com/v2/Http.svc/Speak)  *// authToken= "Bearer " + <剛剛得到的 token 字串>;*  private static void Speak(string authToken){  string strtext = “井民全”;  *// 你想要讓電腦說話的文字*  string string strLanguage="zh-tw"; *// 你想讓電腦說話的語言 (中文)*  *// 等一下要送出去的 Speak API*  string uri = "http://api.microsofttranslator.com/v2/Http.svc/Speak?appId=&text=" +  HttpUtility.UrlEncode(strtext) + "&language=" + strLanguage;    *// 建立 web request, 以便送出 speak 功能*  HttpWebRequest httpWebRequest = (HttpWebRequest)WebRequest.Create(uri);  httpWebRequest.Headers.Add("Authorization", authToken);  *// 送出指令, 並且等待回傳的語音串流*  response = httpWebRequest.GetResponse();    *// 建立媒體播放器, 播放 Server 回傳的合成語音串流*  using (Stream stream = response.GetResponseStream()) {  using (SoundPlayer player = new SoundPlayer(stream)) {  player.PlaySync();  }  }  } *// end of speak method* |

**References**:

1. Text-To-Speech Wikipedia, <http://en.wikipedia.org/wiki/Text-To-Speech>
2. IBM Text-to-Speech Demonstration, <http://www-01.ibm.com/software/pervasive/tech/demos/tts.shtml>
3. Microsoft Translator, <http://www.microsofttranslator.com/>
4. Google TTS, <http://translate.google.com.tw/>
5. 工研院文字轉語音 Web 服務, <http://tts.itri.org.tw/about_us/introduction.php>
6. 查閱 client id & client secret, <https://datamarket.azure.com/developer/applications/>
7. Obtaining an Access Token, <http://msdn.microsoft.com/en-us/library/hh454950.aspx>
8. Speak Method, <http://msdn.microsoft.com/en-us/library/ff512420.aspx>
9. <http://stackoverflow.com/questions/9937421/no-sound-from-client-side-using-microsoft-translator>