

第一章 緒論

1.1 研究背景

二十一世紀，可說是一個資訊時代，一個網路全面起飛的時代。這裡所提到的網路包含三種網路：電信網路、網際網路(Internet)、電視網路。但是由目前的发展來看，未來網際網路會慢慢的把電信網路與電視網路整合在一起，讓人们的生活與網際網路更加的密不可分。本研究與網際網路和電信網路的整合相關，以下簡介了網際網路與電信網路的背景。

網際網路的服務有很多種，其中，全球資訊網(WWW)算是最熱門也最普遍的一種，因為全球資訊網結合了文字、聲音、影像、電腦動畫等多媒體顯示效果，讓使用者可以很容易的在網路上存取各種多媒體資訊。

網頁(Web Page)，就是就是全球資訊網的顯示文件，透過瀏覽器(Browser)，使用者就可以看到網頁要呈現的豐富資訊。在編輯網頁的語言裡最常用的，莫過於 HTML (Hyper Text Markup Language) 了。HTML 是一種網頁呈現語言，是在 1989 年由 Tim Berners-Lee 所發明的，利用它可以讓瀏覽器知道該如何處理及呈現文件的內容。HTML 具有跨平台特性，使用者可以清楚地看到資訊在其豐富多樣的視覺效果下呈現。對於網際網路能夠如此蓬勃發展，HTML 實在功不可沒。但是隨著網際網路的各種應用不斷發展，單純的 HTML 已不能滿足網際網路的所有需求了。

XML(eXtensible Markup Language，可擴充文件語言)是 W3C 在西元 1998 年二月訂定的新一代網頁語言【6】，我們可以將其稱為第二代的網頁語言以別於第一代網頁語言 HTML。XML 比 HTML 更具彈性，更容易添加新功能，但相對地也比 HTML 複雜。XML 至今誕生五年，已經幾乎成了網路上的世界語言，也為下一波網路化的服務奠定根基。XML 會如此熱門是因為具備了以下

優點：

1.開放式平台：XML 與 HTML 最大的不同在於 XML 將資料從一個同時包含視覺呈現與資料內容的網頁文件中抽離出來，所以 XML 文件可以讓文件中包含的資訊，在任何支援 XML 格式的應用程式或平台間傳遞。無論是建立網頁、包裝訊息、文件整合、定義兩家公司交流所使用的語言與字彙等等，都可以有很好的表現。

2.延伸性 (Extensibility)：在 XML 的文件裡，可以根據功能或語意上的需要自行定義標籤，定義良好的標籤可以增加 XML 檔案的可讀性。在各領域、各行業、各種特殊用途上的資料交換，都可以定義一套恰當的 XML 標籤來描述資料內容。

3.結構性 (Structuralism)：XML 資料的結構是以嚴謹的樹狀結構所組成，每個資料單元都由起始標籤與結束標籤包裹起來，不似 HTML 不一定同時有起始標籤與結束標籤那麼的鬆散。此一特性可以讓存在於 XML 文件裡的大量資訊，可以顯示出資料之間清楚的結構性。

XML 因為具有上述的特點，所以運用 XML 剖析器 (XML parser)，透過 DOM (Document Object Model) 技術【21】，我們可以寫程式將資訊從 XML 文件中取出或者放入，如此就可以輕易的操縱 XML 文件資訊，達到應用程式間資料交換的目的，並完成應用程式所要執行的任務。

另外，就電信網路的觀點來說，很久以前，當我們需要向某個機關或單位詢問事情時，我們會親自到該機關的櫃臺去詢問，後來，慢慢有了語音按鍵系統(Interactive Voice Response, IVR)，人們不必親自到場詢問，可以打電話到該機關提供的總機，靠著語音按鍵系統的指示按鍵選擇及轉接，獲得想詢問內容。由於電信網路與網際網路悄悄的在整合，近來又有語音對話系統(VoiceXML

系統)的誕生，讓人們不但可以不必親自到詢問櫃臺，可以用語音與電話總機對話，而取得資訊，更因為這樣的語音對話系統與網際網路整合，所以使用者可以獲得的資訊，比以前傳統式的語音按鍵系統更加的豐富了。圖 1-1 說明了這樣的演進。

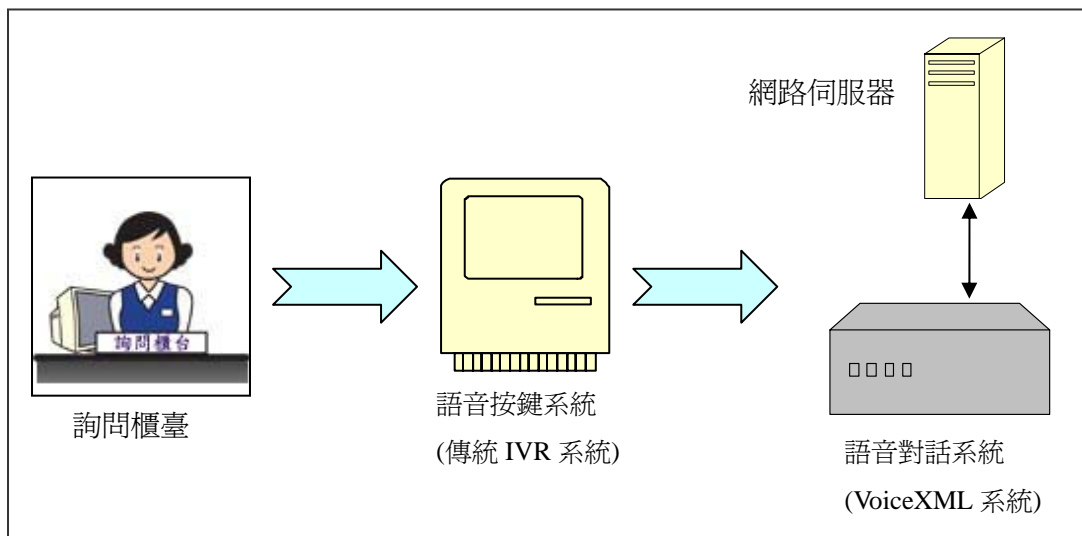


圖 1-1 詢問方式的演進

目前世界科技演進的趨勢，是朝向「全方位(Ubiquitous)網路應用服務」的走向來發展，這個發展趨勢的願景是「Anytime, Anywhere, Anydevice」，也就是希望在任何時間、任何地方、使用任何裝置，都可以順利的上網獲取資訊。如圖 1-2 所示，位於伺服器內的所有資訊，在經過「過濾」與「轉換」的機制之後，就可以傳達給任何裝置，例如個人電腦、個人數位助理(PDA)、電話等等。而本研究的定位就是圖中的「轉換」模組，希望將 HTML 網頁上的資訊經過轉換後，可以讓使用者經由電話輕鬆的獲得資訊。

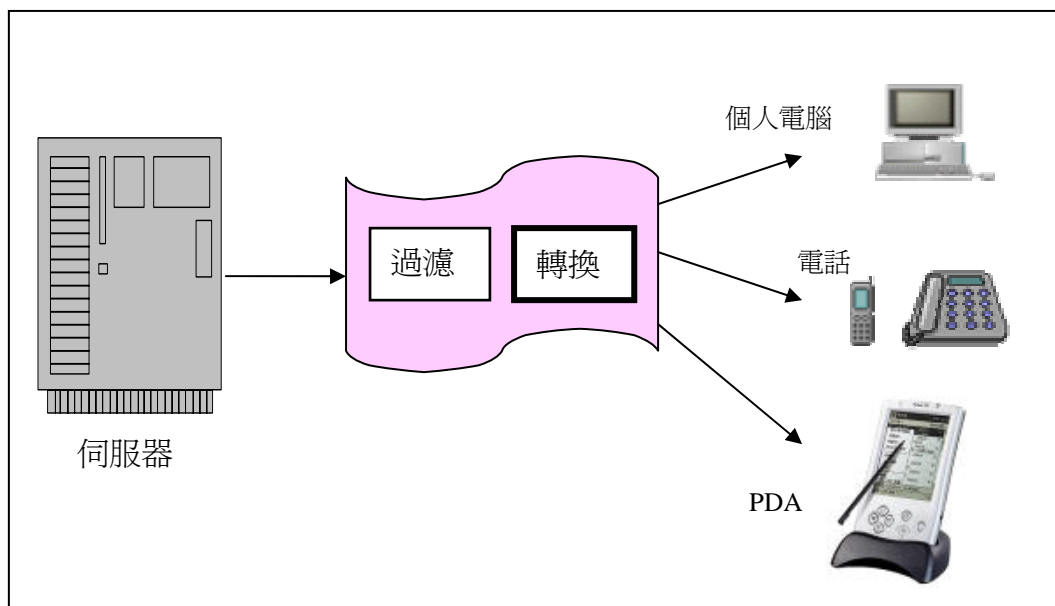


圖 1-2 「全方位應用軟體」概念圖

由於電話網路與網際網路的整合，存取網頁的方式就不會只有「使用個人電腦上網」這樣一種而已，主導全球資訊網發展的主要機構 W3C(World Wide Web Consortium) 目前正制訂多樣化模式存取網站的服務機制【10】，提供使用者除了可以透過視覺瀏覽網頁，未來更可以透過語音辨識模組，利用特定裝置與電腦進行互動對話，進而得知網頁的內容。雖然現在一般電腦可以看到的網頁都是以 HTML 語言所編輯，但是若要可以利用語音來瀏覽網頁，VoiceXML (Voice Extensible Markup Language) 已經是用來設計這種機制的標準了【8】。

VoiceXML 是 XML 的一種延展，代表人與電腦對話的一種標記語言，可讓人們利用語音與電話存取網路上的資訊與服務。如同 HTML 定義網路文字、影像的顯示與傳遞一般，VoiceXML 定義了使用者與網頁間的語音互動行為，使得網頁資訊可以透過電話傳遞。圖 1-3 表示 HTML 網頁使用者與 VoiceXML 語音網頁使用者與網頁伺服器之間互動的情形。瀏覽 HTML 網頁的人，在使用個人電腦上的瀏覽器時，瀏覽器會透過網際網路 Internet 來對網頁伺服器請求瀏覽某個網頁，然後網頁伺服器就會回應，將某個 HTML 檔案交給瀏覽器讓使用者瀏覽；而使用電話來瀏覽網頁的使用者，需要透過公共電話網路(Public

Switched Telephone Network, PSTN)來連上語音伺服器，語音伺服器其實就是電話網路與網際網路的開道器，也是語音網頁的瀏覽器。然後語音伺服器會經由網際網路來連上網頁伺服器，並且請求某一個 VoiceXML 文件，接下來網頁伺服器就會將該 VoiceXML 文件回應給語音伺服器，讓電話使用者可以使用語音互動的方式來瀏覽語音網頁。

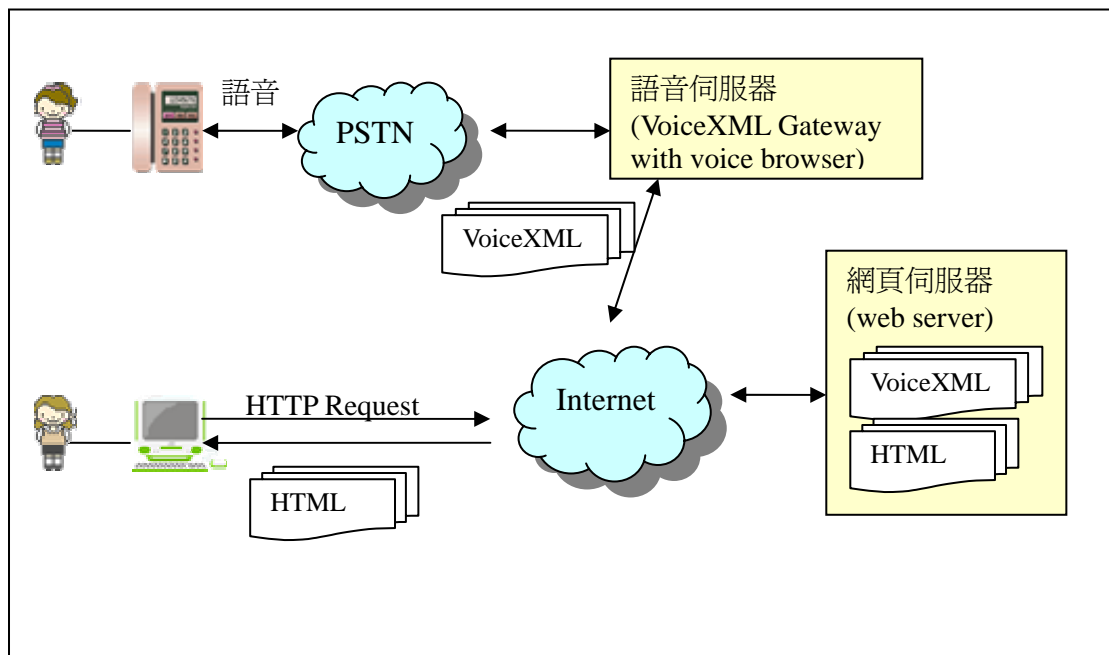


圖 1-3 HTML 使用者與 VoiceXML 使用者示意圖

HTML 網頁需透過網頁瀏覽器來讓人們「看」與「點選」，而 VoiceXML 網頁則需透過語音瀏覽器（Voice Browser）來讓人們「聽」與「說」，兩者都達到了存取網路資訊與服務的目的。在 Voice XML 的架構下，人們可透過語音播放的提示資訊、口述命令、語音或電話按鍵輸入，與 VoiceXML 語音瀏覽器展開互動，然後立即可以「聽」到目前網路上的種種資訊。

表 1-1 是 HTML 應用程式與語音應用程式的比較，HTML 應用程式的使用者是透過網址來存取 HTML 網頁應用程式，而語音應用程式的使用者則是以撥電話號碼的方式來存取語音網頁應用程式；HTML 應用程式是以網頁瀏覽器

向網頁伺服器發出對某個 HTML 頁面的請求，而語音應用程式則是以語音瀏覽器向外部的網頁伺服器發出對某個 VoiceXML 文件的請求；HTML 應用程式是使用網頁瀏覽器來呈現 HTML 視覺網頁，並且使用者可以利用鍵盤與滑鼠來回應此應用程式，而在語音應用程式中，語音瀏覽器將 VoiceXML 網頁以對話和提示的方式來呈現，讓使用者可以透過語音瀏覽器，使用語音或者按鍵來與語音應用程式互動。

HTML 應用程式	語音應用程式
使用者透過 URL 存取頁面	使用者透過撥電話號碼存取應用程式
網頁瀏覽器向網頁伺服器發出對某個 HTML 頁面的 HTTP 請求	語音瀏覽器向外部網頁伺服器發出對某個 VoiceXML 文件的 HTTP 請求
網頁瀏覽器呈現 HTML 以建立可視網頁，該頁面回應鍵盤和滑鼠輸入	語音瀏覽器可以透過語音或按鍵輸入，將 VoiceXML 表現為對話和提示

表 1-1 HTML 應用程式與語音應用程式的比較

1.2 研究動機

一直以來，人們都是以視覺在瀏覽網頁並接收資訊，例如在網路上查詢考試榜單、在網路上瀏覽每日新聞、看餐廳或電影介紹、查詢股票報價等等。但是人們並不是無時無刻都在電腦旁邊，當我們不在辦公室、不在家裡，卻又要接收資訊時，電話就扮演了一個很方便很普及的溝通角色了。隨著 3G 手機時代的來臨，目前手機業者也已深切體認電話帶給大眾生活的便利性，所以現在做在手機上的服務也越來越多，電腦已經不是上網瀏覽網頁、存取網路服務的唯一工具了。

除了一般人需要透過電話來獲取資訊外，視覺障礙者因為無法看到網頁，所以更需要使用電話來瀏覽網頁。本研究想從聽覺的角度來考量，希望在沒有視覺的幫助下，光是靠聽與說，亦可以獲得原本應該以視覺得到的網頁資訊。在 HTML 網頁裡，表格內容一向是一定要親眼看到，才能夠清楚了解的資訊。但是網頁上的表格對於身心障礙者所使用之螢幕閱讀機或其他念網頁的軟體

來說，實在是一個挑戰：應該怎麼念，聽的人才會了解呢？

本研究試著從表格資訊這個點切入，研究並分析歸類網頁上的表格類型，根據每個不同的類型設計不同的對話模式，讓使用者可以藉由電話與語音平台對話互動，希望在一般網站上能夠看到的表格資訊，也可以在語音瀏覽器上以語音網站的方式來呈現給使用者。對於一些視覺障礙者來說，用眼睛瀏覽網頁並非容易的事情，本研究可以幫助他們不用視覺，用聽覺亦可以得到本來需用視覺才能得到的網頁表格資訊。

1.3 研究目的

本論文的研究目的有以下幾點：

- (1)分析 HTML 網頁表格的類型。
- (2)探討如何設計 VoiceXML 與使用者互動的對話模式，才能夠讓使用者了解原來 HTML 所包含的表格資訊內容。
- (3)將一般瀏覽器能呈現的含有表格內容的 HTML 網頁，轉換成語音瀏覽器能夠呈現的版本。
- (4)探討如何將好幾個 HTML 表格網頁，可以方便的合併成一個語音網站，且此語音網站的結構，可以讓使用者自由設計。

如圖 1-4 所示，傳統的語音按鍵系統後來慢慢的進步成語音對話系統，而本研究就是架構於語音對話系統(VoiceXML 系統)之上，希望能對語音對話系統的應用有所貢獻。

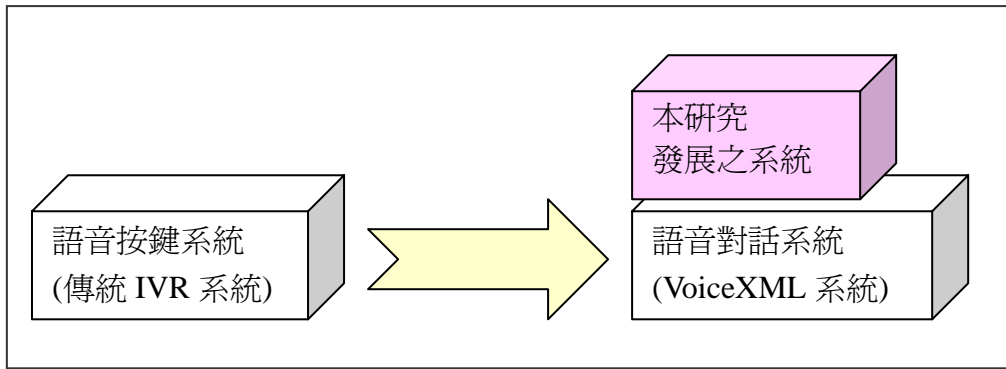


圖 1-4 本研究定位圖

1.4 論文架構

本論文共分為五個章節，第一章為緒論，介紹研究背景、研究動機、研究目的與論文架構；第二章為相關文獻探討，針對多樣化模式存取網站的服務機制、XML 的語音技術—VoiceXML，與現今語音瀏覽網頁的實例來做探討；第三章為網頁表格資訊自動對話模式設計，提出電話語音網站製作中心、網頁型表格資訊對話模式、資料型表格資訊對話模式的設計方法與細節；第四章是系統架構與實作，介紹系統架構與實作環境；第五章為結論與未來發展，提出各個語音系統之間的分析比較，與探討未來的發展方向。