

本期要目

- | | |
|---------------------------|---------|
| 壹. ROCLING XVI-專題演講簡介 | 第二~三頁 |
| 貳. 專文-多媒體人機通訊實驗室-吳宗憲, 簡仁宗 | 第四~十五頁 |
| 參. 專文-自然語言處理與圖書資訊學-陳光華 | 第十六~二十頁 |

ROCLING-XVI開始報名

「第十六屆自然語言與語音處理研討會」謹訂於 9/2 9/3 假台北縣福華翡翠灣舉行, 本次會議共收錄 38 篇論文, 另特別邀請微軟亞洲研究院的副院長 Dr. Hsiao-Wuen Hon 及 Dr. Eric Chang 發表兩場專題演講, 介紹微軟中文處理研究及相關產品發展, 另外, 大會還將安排 (1) Industry Panel: 邀請業界朋友介紹相關產業發展 (2) 智慧型 Web 技術研討會、(3) 學生論文研討會等活動。

本次會議一律採線上報名, 凡為本會會員報名者皆享有優惠價, 會員身份失效者(本年度未繳費者)請繳交會費方可享有優惠, 非會員可同時加入會員亦享有優惠價。報名截止日為 8/15 日, 8/16 日以後報名者加收 300 元, 8/27 日以後將關閉線上報名, 擬報名者請至大會現場報名, 現場報名將加收 500 元。

大會為了服務與會者, 特別安排「交通」及「代訂住宿」之服務, 交通方面將安排中研院⇔會場、台灣大學⇔會場及會場⇔台北火車站的接駁車, 住宿方面則提供中研院活動中心及福華翡翠灣住房代訂服務, 上述兩

項服務請於註冊時依個人所需勾選之。

註冊網站:

<http://www.iis.sinica.edu.tw/Conference/ROCLING04/>

會員大會開會通知

第八屆第二次會員大會謹訂於 9/2 日假台北縣福華翡翠灣舉行, 本次會議除了例行之年度會務工作報告外, 並將頒發本屆碩士論文獎。開會通知將於八月中旬郵寄予本會會員, 通知書內將夾附「委託書」, 若確定無法出席會員大會之會員者, 請簽妥「委託書」後傳真回本會, 以利大會順利進行。

六月份學生出席國際會議補助

補助會議: COLING-2004

論文題目: Automated Alignment and Extraction of Bilingual Ontology for Cross-Language Domain-Specific Applications

補助學生: 葉瑞峰 (成功大學資訊工程系)

補助金額: 美金 1,000 元

ROCLING XVI

Keynote Speeches

Why should people care about Computational Linguistics and Speech Processing technologies?

Hsiao-Wuen Hon

Assistant Managing Director, Microsoft Research Asia (MSRA)

Abstract:

Should a student choose Computational Linguistics and Speech Processing as his major instead of semiconductor? Why do companies like Microsoft spend billions of dollars in R&D of Computational Linguistics and Speech Processing? In this talk, I will provide a historical aspect of why Microsoft started R&D in Computational Linguistics and Speech Processing, how these technologies got deployed into mainstream

Microsoft products and finally what products and features you can expect to be benefited from continuing development of these technologies. There are unbounded applications and potential benefit Computational Linguistics and Speech Processing technologies can bring to millions of users.

Biography:

Dr. Hsiao-Wuen Hon is the assistant managing director of Microsoft Research Asia (MSRA) . He is supervising research in speech, natural language processing, information retrieval, internet search, audio/video indexing retrieval and other related areas. Before joining MSRA, Dr. Hon was the Architect in Speech.Net at Microsoft Corporation. Besides overseeing all architectural and technical aspects of the award winning MicrosoftR Speech Server product (VSLive! Editor Choice Award), Natural User Interface Platform and Microsoft Assistance Platform, he is also responsible for managing and delivering statistical learning technologies for Natural Interactive Service Division (NISD). Dr. Hon joined Microsoft Research as a senior researcher at 1991 and has been a key contributor of Microsoft's SAPI and speech engine technologies. Before joining Microsoft, Dr. Hon worked at Apple Computer, Inc., where he was a principal researcher and technology supervisor at Apple-ISS research center. Dr. Hon led the research and development for Apple's' Chinese Dictation Kit, which receives excellent reviews from many industrial publications and a handful of rewards, including Comdex Asia'96 Best software product medal, Comdex Asia'96 Best of the Best medal and Singapore National Technology award.

Dr. Hon received the B.S. in Electrical Engineering from National Taiwan University; and M.S. & PhD degrees in Computer Science from Carnegie Mellon University. While at CMU, Dr. Hon is the co-inventor of CMU SPHINX system on which many commercial speech recognition systems are based on, including Microsoft and Apple. Dr. Hon is an international recognized speech technologist

and has published more than 90 technical papers in various international journals and conferences. He is currently a senior member of IEEE and an associated editor for IEEE Transaction of Speech and Audio Processing. He has also been serving as chairs and reviewers for many international conferences and journals. Dr. Hon holds 35 US patents.

From the Lab to Ubiquity: Speech Technology's Road to Mainstream

Eric Chang

Assistant Managing Director, Microsoft Research Asia Advanced Technology Center

Abstract:

Invention, development, and widespread adoption of any discontinuous innovation follow a technology adoption life cycle, as discussed by Geoff Moore in his book "Crossing the Chasm". Speech technologies have been researched in the lab since the early 1900's. In this talk, I will provide a brief history of the development of speech synthesis technology and speech recognition technology and offer my view on where these technologies are on the technology adoption life cycle.

Biography:

Eric Chang joined Microsoft Research Asia in July, 1999. Eric is currently the Assistant Managing Director of MSR Asia Advanced Technology Center, where he is in charge of program management, operations, and technology incubation. Previously, Eric was the research manager of the speech group at MSR Asia. Some results from his group are the Mandarin speech recognition engine included in Office XP and the Mulan bilingual text to speech system. Prior to joining Microsoft Research, Eric was one of the founding members of the Research group at Nuance Communications, a pioneer in natural speech interface software for telecommunication systems. While at Nuance, Eric worked on various projects involving confidence score generation, acoustic modeling, and robust speech detection. He also led the technical effort to develop the Japanese version of the Nuance product. This project led to the world's first deployed Japanese natural language speech recognition system. Eric has also developed speech recognition algorithms at M.I.T. Lincoln Laboratory, invented a new circuit optimization technique at Toshiba ULSI Research Center, and conducted pattern recognition research at General Electric Corporate Research and Development Center.

Eric Chang graduated from M.I.T. with Ph.D., Master and Bachelor degrees, all in the field of electrical engineering and computer science. While at M.I.T., he was inducted into the honorary societies Tau Beta Pi and Sigma Xi. Eric has published papers in the fields of speech recognition, neural networks, and genetic algorithms in various journals and conferences. He is the author of several granted and pending patents. His research interests are spoken language understanding, machine learning, and signal processing.

多媒體人機通訊實驗室

Multimedia Human Machine Communication Lab

吳宗憲 簡仁宗

國立成功大學資訊工程學系

本實驗室近年來致力於『多媒體資訊處理』的研究。隨著電腦硬體效能的提昇以及網際網路的蓬勃發展，多媒體資訊如雨後春筍般出現。為了有效率的自動化處理龐大繁雜的多媒體資訊，實驗室在數年前即針對文字、聲音及影像視訊等多媒體資料進行相關的研究，發展出具實用性質之技術如：語音合成、語音及語者辨識、人臉辨識、語音文件檢索以及多模式人機介面等多媒體技術。除了應用技術的開發外，本實驗室亦著重於基礎理論的研究。於語音參數調整技術、特徵參數分析及模組化、統計式機率模組分析及相關理論的研究多年，並輔以實作來驗證，均有不錯的成果。

國立成功大學多媒體人機通訊實驗室於民國八十八年成立，由吳宗憲教授及簡仁宗副教授共同主持。實驗室前身為王駿發教授與吳宗憲教授所主持的『中文資訊處理實驗室』，本實驗室是兼顧學術性之研究以及實務應用系統之開發。指導教授均為語音/語言資訊系統設計的專家，歷年來在數位語音訊號處理、口述語言處理的理論與實務均有深入之研究心得，擁有豐富的實作經驗。另外，在跨院校、跨領域整合、臨床醫學及工程整合之計畫執行具相當多之經驗與成就。近年來，更致力於手語溝通輔具、人機介面、語言知識表達處理以及人臉辨識和資訊檢索的研究。在實驗室同學的共同努力下，有相當豐碩的成績，成果發表於國際期刊和學術會議中。除了論文發表外，各項系統之實作並多次於國內外貴賓參觀時，對外作實際展示。

本實驗室致力於下列各領域方向之研究與發展：

語音處理基本理論和技術

本研究探討語者特性之語音調整，包含語音特徵參數分析，模組化以及語音參數調整技術語音參數調整技術上行生的問題。

語言知識表達研究與應用發展

本研究不僅探討中文語言處理的基本問題，而且也將研究語音參數調整技術成果擴及手語轉譯與生成處理。

資訊檢索研究

本研究主要在探討多媒體資訊的分類、檢索、和擷取等問題。

生物特徵資訊處理技術

本研究主要在探討在複雜背景下，執行人臉的偵測、追蹤及辨識的人臉辨識系統，進而擴展到生物認證融合技術。

目前實驗室從事的研究課題包括：

- (一) 手語溝通輔具
- (二) 電腦語音處理
- (三) 多模式人機介面
- (四) 自然語言處理
- (五) 人臉辨識
- (六) 多媒體資訊檢索

(一) 手語溝通輔具

聲音或語言機能喪失的聽語障礙者，常常發生難以與一般人正常溝通或溝通時發生明顯的障礙，因此遭遇到就醫、就學、就業等方面極大阻礙，而無法融入社會的主流。現年來，在全球化之電子資訊、網際網路、行動通訊等科技產業的發達與成熟，國內各界也在此一產業範疇內佔有顯著性的地位，然而卻未能妥善的應用語音資訊與輔助科技在特殊教育與身心障礙者醫療輔助系統的研究發展，而近年來相關的本土化輔助系統尚屬研發的萌芽階段。有鑑於此，本研究現階段發展方向乃經由語意理解擷取情緒資訊在提供台灣手語語言學之分析，以作為手語溝通輔具設計之理論根據，並且協助教學訓練成果之評估。

本研究目的在應用自然語言處理、中文語音訊號處理技術及本土聽語障礙族群實際溝通輔助需求，研發出本土化 PC-based 台灣手語轉語音溝通輔具，以改善其日常生活中的溝通表達。其設計原理，乃以中研院自動斷詞程式及知網 (How-Net) 為基礎，以句型樣版及概念從屬之語格文法 (Case Grammar) 為構句基本架構，透過語法剖析、語意分析以及虛詞補綴等自然語言技術，建立關鍵詞彙預測完整文句之轉譯系統，以提供使用者經由輸入簡單關鍵詞彙，即可獲得完整文句及合成語音。

近年來，本實驗室有感於國內聽語障口、手語相關研究的不足，因而決心投入聽語障相關之科技輔具研發及基礎建設的革命性工作，亦獲得許多具體成果：

- (1) 榮獲 89 年度身心障礙輔具創作競賽第一名
- (2) 90 年度教育部軟體創作競賽優等獎
- (3) 數項專利申請與產學合作案

並且將研發之成果，具體投入於聽語障教學訓練之實務工作，其中個案應用已轉介並推廣於台南啟聰學校國中、小部、南師特教系教學實務、國科會南區輔具中心、內政部聽語障研發中心等單位。

本研究主軸乃根據 AAC 溝通輔具設計的理念，其架構包含了視覺回饋之虛擬鍵盤/人物及語言集 (訓練為標的之教材)，透過安置國語口手語、手語手勢碼為基礎的虛擬鍵盤，依手語符號所對應之字、詞 (語言集)，運用統計匹配的檢索技術選取適當之手語符號，並且進而

經過自然語言處理將此輸入字串轉譯成合乎語法及語意的文句，或是由中文文句產生手語及語音的輸出；另一方面透過語音辨識及多變量樣式比對技術來分析一般及聽語障語者的語音，結合測試語彙庫及聲道模型的建立，進一步提供臨床及教育進行構音及語彙能力測驗的應用；最後，將發展出智慧型口手語溝通輔助暨教學代理人系統，此一虛擬人物具聽(構音評量、詞彙難易度評量、語音辨識)、說(唇/手語合成、語音矯正與合成)、讀、寫(文字、手語、語音、唇語互轉譯)的能力。

(二) 語音處理

語音辨識的功能在於將輸入的語音信號轉為文字輸出。語音辨識首先藉由大量的語料，訓練出語音聲學模型 (Acoustic Model) 及語言模型 (Language Model)，這些模型設計依產品應用需求，可設計出關鍵詞辨認的語音辨識引擎或是大字彙連續語音辨識引擎語音處理部分。主要相關研究課題如下：

A. 中文文句翻語音系統

本實驗室所發展之中文文句翻語音系統 (TTS) 系統，在語音輸出部份已發展到無限量文句系統，一般稱此類系統為文句翻語音(Text-to-Speech: TTS)系統，雖然此類系統複雜度高，然而可合成任意文句，應用的範圍極為廣泛。而個人數位助理系統即為一架構在 TTS 系統上的應用軟體。

TTS 系統大致上可分成三個部份：文句分析、音韻訊息 (Prosodic Information) 產生、語音合成。當系統在文句輸入後，隨即分析文句並配合音韻訊息產生器產生文句的音韻特徵，如各音之音高(pitch)、音量(energy)、音長(duration)、停頓(pause)及句調(intonation)等，再載入相對應的語音合成單元，運用訊號處理技術調整語音合成單元，然後輸出語音。

在文句分析處理的步驟方面，包括斷辭處理及破音字的問題處理，機器要如何去判斷一個字到底該發何種音也是在我們文字分析模組處理上的主要問題，而本系統往往會根據前後的字根與這個字的辭意去調整他的發音系統。

而一套好的 TTS 系統，除了要有好的合成音質外，也要能合成具有抑揚頓挫的句子。本實驗室發展之 TTS 系統使用具有聲調的單音節作為基本合成單元，而這些合成單元是由一個大的語音資料庫中選取的。其流程包括基週偵測及平滑化、語音單元的平滑化、頻譜參數的求取、合成單元挑選、以及人工檢驗。而音韻產生的規則，主要包括聲調調整、句調調整、音長調整、聲音大小調整、及停頓點以及停頓時間長短決定。

B. 強健性語音辨認

在語音辨認系統由實驗室走向實際應用的過程中，不同語者和環境產生的變異性往往造成辨認的正確率大幅下降，因此系統的強健性(Robustness)成為不可或缺的一環。其中包括了語者不特定的處理、背景雜訊的消除以及通道雜訊的補償。在過去的研究中，我們也開發一種具有非監督及累進學習能力之語音辨識技術，它的特色是在人與電腦溝通的同時能夠漸進且不需要知道語者講話內容的情況下將原始語音模型調整到最新環境，因此，不需要任何的註冊及訓練即可累進提昇語音辨識的效果。相關理論已刊載在著名的 *IEEE Transactions on Speech and Audio Processing*，並且有數篇論文發表在語音處理領域重要期刊及會議。

C. 遠距離麥克風陣列語音訊號加強

麥克風陣列語音處理技術可以偵測遠距離語者發話的方向，進而計算聲波至麥克風間的時間延遲，經過時間延遲的補償後可以得到加強過的語音訊號。在先前的研究中結合了模型參數調整的技術，以消除遠距離測試語音與語音模型參數間的不匹配，讓使用者可以透過遠距離免持麥克風方式與電腦進行溝通，這種技術在國內外相當具有前瞻性。它是一套非特定語者之語音辨識系統，可以辨識中文連續數字。

D. 語者驗證

語者驗證是用來對使用者進行身份查驗之用，目的在於確認使用者之身份以載入該使用者在資訊查詢上的偏好，如此可有助於漸近式地提昇系統所回覆之資訊，其直接可用性及其讓使用者無須再另外進行一次篩選的動作，可以大大減少使用者在查詢上的挫折感而提升系統之親和性。本研究主要是以在背景介紹中所提及的文字提示為主的語者驗證架構來進行。在以文字提示為主的語者驗證架構中，其語音模型由隱藏式馬可夫模型表示，故而在系統建構開始，需先訓練出一套非語者相關之語音模型。接下來，當使用者為首次使用系統時，系統亂數產生數道文字序列，由使用者錄製，用於將非語者相關之語音模型調適到所謂的語者相關之語音模型，才能用於接下來之驗證流程中。

E. 語者說話速度調適

除了環境和語者之間的差異性之外，語者本身的差異性，也是影響辨識系統效能的因素之一，尤其是語者說話的速度。由於目前的語音辨識系統無法解決語者速度的問題，所以當使用者在使用語音辨識系統時，常常要被迫使用適合系統的速度，通常是刻意放慢速度來說話，對使用者而言並不方便，也不自然。所以如何有效克服說話速度的問題，使得原本以普通速度語料所訓練出來的辨識系統，在辨識快速語料如新聞或是廣播，達到較高的辨識效能，是本研究室的另一研究重心。除了利用說話速度來分析語音特性之間的關係，使用模型化的方法來描述語音段長度的音長模型，也是對於說話速度方面的一個研究重點。

音長模型的研究重心多著重於如何的建立音長模型或是加入音長模型的隱藏式馬可夫模型，以及音長模型的應用。因此，如何精確、有效地描述語料速度，使用何種統計分佈作為語音之音長模型是最適當的與如何進一步根據線上所收集到的少量語料將音長模型調適到更適於目前的測試環境，是我們目前研究中語音項目的重點。關於語音模型使用音長模型調適的部份，我們提出以近似貝氏估測理論為主的漸進式調適方式。此調適方法我們將針對以高斯分佈與波式分佈為主的音長模型參數進行調適。此具可調適性之音長模型將整合到以隱藏式馬可夫模型為主的語音辨識系統中。

(三) 多模式人機介面

多模式互動溝通(Multi-modal/Multimedia Interaction)環境建立已是科技發展與應用層面中不可或缺的趨勢；因此，在人機溝通介面上，除了結合現階段語音辨識/合成技術、網際網路多媒體檢索及無線網路通訊技術來提供使用者最便利的操作模式，另一方面，具高度親和力及智慧型互動的影音輸出入介面亦越來越受到重視。現年來，由於半導體製成、電子、電機及資訊等科技的蓬勃發展與成熟，相關科技的應用亦受到業界相當的重視，使用者於選購階段，

已不僅僅考量功能與操作性能，舒適、人性化的操控環境更為主要的依據，其中，視訊、影音、通訊、智慧型等附加價值的提供成為各家廠商訴求的重點，深具情緒反應、理解及自動對話特性之產品，更為各界積極投入的研發領域。另外亦需考量傳輸頻寬、傳輸效能、壓縮編碼等技術，此類關鍵技術的突破，可進一步地提供系統廠商整合更具效率、人性化的智慧型環境。本研究的目的是在於整合語言學以及情緒心裡學等各學科的技術，局部理解文句中的語意內涵，並擷取出純文字的自然語言對話中發話者的情緒表現及情緒轉折。先前的研究方向著重在經由語意理解擷取情緒資訊結合自動語音合成、自動唇型及上半身動作合成、使用者自訂介面等技術，達成一應用於對話型行動資訊系統中之虛擬人物動畫介面。近年來，主要發展方向包括：a. 以情緒為主的文句剖析器、基於詞彙之情緒參數定義和擷取以及適用於少量資料之參數分類。b. 建立一個擬真情緒反應之情緒代理人系統，系統包含情緒輸入、處理及輸出三大模組。其建構完成之系統可應用各式查詢、服務、看護以及一般娛樂等系統。情緒理解之所以困難，主要是由於幾項重要因素：

- (1) 自然語言的多樣性
- (2) 知識表示集的不確定
- (3) 心理認知過程無程式化

A. TTVS 系統

TTVS(Text-to-Visual-Speech)系統是將對話系統中的文字回應配合文字轉語音系統(TTS)，在顯示器上以虛擬人形輸出，並以自然的表情與使用者互動。TTVS 系統包含唇型動畫、表情生成及手勢動畫三部分。分述如下：

- (1) 唇型動畫模組經由預先定義的唇參數變化多列，將輸入的文字轉換為虛擬人頭上的網格點變化。首先經由一組事先拍攝的唇型變化影片擷取出唇形變化參數。再由一預先建構好的立體人頭網格模型，將所擷取的情緒變化轉換為表情參數，並以動態人頭表情變化配合嘴型變動來回應出文字及情緒變化。在最後輸出時為求整體畫面能一致，因此必需考慮語音及影像輸出的同步問題。
- (2) 表情生成模組則經由表情控制點的移動形成臉部網格變化，並重疊於臉部網格模型上，形成具表情的唇型動畫。人的表情基本上可以獨立為數個部分，如嘴角的位置，眉毛的高低，眼睛的大小等，FACS 系統將臉部分為四十四個可參數化的獨立單元，稱為 Action Unit (AU)。經由拍攝不同情緒之臉部影像，以人工或是影像處理方式求得臉部各單元的運動參數，並應用類神經網路之倒傳遞演算法求出情緒參數與臉部單元的對映。
- (3) 手勢動畫模組則是以階層式的手勢描述語言控制肢體骨架系統，並同步影響包覆其上的網格模型，形成手勢動畫。綜合以上三者，可輸出為一具情緒反應的虛擬人物動畫。相較於前述的臉部運動，手勢運動有著根本上的不同。臉部運動除了下巴轉動之外，其餘都是由臉部肌肉互相牽引、變型等非剛體運動所引起。手勢運動則是由骨骼間的轉動、平移等剛體運動所造成。為達到手勢運動合成，本計劃預計建立一擬生理結構之手部模型，包含手部骨架系統及手部網格模型。再經由手勢運動描述系統來合成手勢。

B. 語音情緒擷取

針對機器視覺、機器人、人工智慧、人機介面及人類互動反應的研發領域，包含情緒偵測、情緒/意圖理解、緊急情況偵測與適度性系統回應等處理，經由心理學的研究得知，情緒的回饋及適當反應對於人類與外界溝通有著不可或缺的重要性，但長久以來在人機介面設計領域上，對情緒的處理還未見到成熟的應用。另一方面，儘管電腦語音辨識、合成技術已有相當之水準，然而合成之電腦語音並無法帶有情緒之反應。主要的瓶頸在於以下數個原因：

- (1) 對於情緒產生的心理及生理機制未完全了解。由於人類對情緒的理解是建立在對事實、環境、及本身特質等項目的綜合且複雜的判斷上，因此，對於真正的情緒產生過程尚未能有一個具指標性且量化的模型。而這點嚴重影響到電腦對情緒的處理。
- (2) 應用層面尚未開發。目前所見到的電腦應用系統，多著重於應用電腦強大及快速的運算能力。因此所開發的系統也是以解決人力無法達成的大量運算及精確度為目標。
- (3) 電腦運算能力受限。雖然有精確而快速的運算能力，對於如情緒這種難以量化的特質，電腦依然束手無策。因此，也侷限了在情緒處理上的發展。
- (4) 對於語音之所以帶有情緒反應並未完全了解。人類將情緒反應在語音上的模型十分複雜，因此，難以精準量化情緒之語音參數，使電腦合成之語音能夠帶有跟人類一樣之情緒。而這點正是情緒化電腦語音系統之主要核心。
- (5) 語音合成技術之限制。目前致力於電腦語音合成之學者，著重的是使合成之語音能夠更清晰平滑，然而此類技術卻限制了情緒化電腦語音之合成的可行性，主要的原因在於人類將情緒反應在語音時，經常是很極端的情況，傳統的語音合成技術無法達成大幅度參數調整的目的。

為了能達成情緒化電腦語音辨識及合成的長遠目標，本實驗室分析影響語音之情緒反應的主要參數，進一步擷取情緒相對應之語音參數型態，用以建立情緒偵測辨識機制、情緒/意圖理解及具情緒參數調整之任意文字語音合成系統，達成人性化溝通之目標。

(四) 自然語言處理

本小組之研究領域主要為中文相關自然語言處理之基本理論探討，與自然語言在實務人機界面應用上之研究。在過去的研究上提出了利用類片語單元(Phrase Like Unit)來解決關於集外詞(OOV)，並配合工研院委外研究計畫的執行提出一套半自動之知識概念模型建構方法，憑藉對相關階層化知識庫資源如 HowNet、WordNet 之熟悉，將之運用於網際網路 FAQ 答問集之系統實現，因其效果卓著該系統之論文獲頒二 0 0 二年度龍騰知識經濟論文獎資訊科技類論文獎與資訊學會論文獎，並將前述知識概念模型建構方法實現於醫療領域之跨語言知識表示。並以此知識表示法建構出具知識推論能力之資訊擷取系統，如網頁採礦、FAQ 答問集擷取之系統。近兩年並將該知識概念模型結合先前發展之口述語言對話系統整合發展出一套可以提供醫療衛教文件查詢，診療科別推論與對話服務之口述語言系統（此一系統之學位論文於二 0 0 三年度獲頒龍騰知識經濟論文獎資訊科技類優等獎，此一系統根據網際網路可連接至成大醫院之網際網路獻上掛號擷取相關掛號資訊，並利用口述語言處理技術在對話中辨認語者之語意動作(Speech Act)並以語意相關規則來記錄對話歷程，並加以推論，以對話者輸入做語意層次之初步理解，令使用者在最習慣之互動方式完成掛號等相關服務。尤其系統中針對成大醫院之門診科別，參酌專家與醫師意見所建構出一千餘條關於疾病症狀與科別之推論規則，

可以有效地減少使用者因對醫學知識之不足，所產生掛錯門診科別之醫療資源浪費，與延誤就診時機之不良現象。

(五) 人臉辨識

近年來，很多研究單位都已經開始探討利用人體的一些生物特徵來做為辨識個人身份的方法，例如人臉辨識、聲紋認證、眼球虹膜比對、指紋或掌紋比對等等。在這些方法中，如果系統只能從一些影像媒體(如影片或照片)取得人臉影像而非實際有受試者接受身分認證時，就只能選擇採用人臉辨識。雖然指紋或虹膜掃描等生物特徵也常被應用於身份識別或驗證之依據，但由於人臉辨識屬於非接觸式的機制，理想狀況下被辨識的目標無須刻意配合辨識系統之動作要求，即可達成辨識之目的，因此可以預期的是以人臉作為身分識別基礎的工具將會被廣泛的應用。

本實驗室針對人臉辨識中常見的幾個問題做深入的探討及研究：

A. 人臉特徵擷取

本研究提出應用二維離散小波轉換於人臉特徵擷取來降低辨識影像維度，取得其低頻的部分並拉成向量形式來當作特徵的方式來取代傳統的特徵臉方式，有效的減少運算的時間與複雜度。擷取出來的特徵影像或特徵值便可以輸入辨識系統來做為比對辨識該影像身分的資料。

B. 人臉模型訓練

主要目的著重在於人臉模型訓練時所用線性鑑別式分析理論的問題改進，並搭配上人臉偵測與最接近特徵點的分類技術，達到良好的人臉辨識效果。我們採用資料分析中常用的線性鑑別式分析(Linear Discriminant Analysis, LDA)方法，在很多文獻上面說明了這個方法對資料維度的縮減跟保持類別間的鑑別性有很好的表現。但是這個方法要特別注意當訓練的樣本數不足時會使類別內散佈矩陣成為非滿秩矩陣（奇異矩陣，singular matrix）而使得 LDA 無法實現的危險，針對這個問題我們也會在系統中提出一個非奇異矩陣轉換的方法來獲得解決。

C. 人臉偵測

如何能在一般影像檔案中確實偵測出人臉位置及大小，並將其正確取出以執行辨識，是人臉辨識系統中必要及必須優先討論的課題。首先，運用機率模型的方式去篩選出影像中可能為人臉的區塊。第二階段是藉由考慮人臉上眼睛輪廓的幾何關係來判斷上個階段所篩選出的區塊是否含有人臉在內。為了取出區塊之內的輪廓，我們採用離散小波轉換後的高頻部分相加，再經過二元化的步驟來取出輪廓。最後，再套用一些眼睛輪廓之幾何關係的規則來判斷是否有成對的眼睛輪廓在區塊內。如果有這樣的成對輪廓，則判斷此區塊內有人臉影像在內。

(六) 多媒體資訊檢索

本研究主題旨在解決多媒體資料文件(例如新聞、廣播)音訊內容之前處理問題。尤其我們已身處多媒體資料量暴增的時代，人工處理、標記、轉譯多媒體資料結構及內容已是耗時耗力的方式。倘若能利用電腦自動地處理音訊資料及轉譯、摘要語音文件，便能廣泛運用於多媒體資料管理、多媒體資訊檢索等重要應用。主要研究課題如下：

A. 新聞音訊切割暨分類

此研究課題旨在自動化地針對新聞音訊文件切割出不同音源、語者、背景之音段，並且將其分類，如此便能分析出新聞音訊之初步流程。接著分析新聞音訊之詳細流程，定出主播、外景、音樂等音段，未來即可運用不同之語音辨識器轉譯內容並提升整體辨識效果。

B. 新聞主題偵測暨切割

待得知新聞音訊詳細流程後，吾人運用語音辨識器轉譯各則報導內容，並自動化地切割暨偵測各則報導之主題。由於新聞主題之變化非常迅速，因此如何建構一個可快速調適且精確分類的主題分類器是主要的研究課題。

C. 語音文件自動摘要

語音是溝通最直接的方式，對於現今所存在的大量豐富語音資訊，需要有效的利用和開發。自動語音文件摘要是很實用的技術，方便人們對於資訊的檢索、瀏覽以及紀錄等應用。一個好的語音摘要系統，也就是做到模仿人類聽覺理解系統，從聲音擷取、語音辨識、語意分析、語意理解進而到語音摘要，一連串的流程環環相扣。也因此目前自動語音摘要存在有幾個問題，首先是語音辨識可能發生辨識錯誤，然後是如何對摘要語音，考慮其重要資訊的萃取保留、語法結構上的合理性、以及語意意涵的關聯性等因素。最後為了讓語音重現，依照摘要單元自動切割合成摘要語音輸出。

如何才能從語音文件中萃取出重要的詞句，並且能夠代表文章意函的內容，這是自動語音文件摘要所要思考的問題。再者，我們並不能只是單純的將這些抽取出來的資訊，好比像是名詞、動詞或形容詞等，做任意的排列組合，這樣所呈現出來的資訊並不具有意義。因此，摘要的另外一個重要的目的，也就是需要探討如何才能生成一則具有意義、合乎文法規則的文章。

自動語音文件摘要的概念，主要從語音聲學(acoustics)、語言學(linguistic)，句法(syntax)和語意(semantics)等方向去解決可能面對的問題。應用特徵參數表示方法，一篇摘要文章，可以被歸納分析成五個要素，包含有：

- (1) 語音辨識信賴度量測
- (2) 關鍵詞的選取
- (3) 語言學知識
- (4) 語意相依法則
- (5) 機率式文法規則

D. 以語音互動的資訊檢索技術

以語音輸入方式進行資訊檢索可以提供相當便利的人機介面，以及資訊較為充足的查詢，特別是對電腦輸入不易的中文。但是這樣的語音檢索技術必須能辨認相當大量且變動的詞彙，檢索系統也必須具備容錯能力。語音技術除了提供較便利的輸入方式，更重要透過語音可以與資訊檢索系統有更好的互動(Speech Interaction)。在檢索過程中許多檢索動作的下達，以語音的方式進行遠比使用滑鼠點選或者輸入指令更為自然有效。這樣的語音互動檢索技術可以顯著提升資訊檢索系統的人機介面能力，開啟相當寬闊研究空間。但是在技術上必須提升現

行系統對的資訊內涵的分析與認知能力。

E. 新聞文件資訊檢索

資訊檢索系統中，文件與查詢問句所具有的共同概念可藉由詞彙顯現，而詞彙本身的語意關係（如關聯詞、狹義詞及廣義詞）亦在概念空間內相互連結。因此，詞彙不僅表現出文件的內容，同時也表達出檢索者的資訊需求。此研究課題主要是運用各種不同機率模型於文件檢索，以其獲得較好的檢索結果。其方法如混和式機率模型、潛在式語意分析以及自動相關回饋等技術。語言模型技術是多數自然語言處理應用，如語音辨認，機器翻譯，文字辨識中不可或缺的關鍵性技術。

實驗室相關文獻

(一) 手語溝通輔具

Chung-Hsien Wu, Yu-Hsien Chiu and Kung-Wei Cheng, “Error-Tolerant Sign Retrieval Using Visual Features and Maximum A Posteriori Estimation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26(4), pp. 495-508, 2004.

Chung-Hsien Wu, Yu-Hsien Chiu and Chi-Shiang Guo, “Text Generation from Taiwanese Sign Language Using a PST-based Language Model for Augmentative Communication,” to appear in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2004.

Chung-Hsien Wu, Yu-Hsien Chiu and Kung-Wei Cheng, “Multi-Modal Sign Icon Retrieval for Augmentative Communication,” *Springer-Verlag Lecture Notes in Computer Science: Advanced in Multimedia Information Processing*, pp. 598-605, 2001.

Chung-Hsien Wu, Yu-Hsien Chiu, and Kung-Wei Cheng, “Multi-Modal Sign Icon Retrieval for Augmentative Communication,” *Lecture Notes in Computer Science Series, Springer-Verlag*, 2001.

(二) 語音處理

Chung-Hsien Wu and Gwo-Lang Yan, “Speech Act Modeling and Verification of Spontaneous Speech with Disfluency in a Spoken Dialogue System,” Accepted by *IEEE Trans. Speech and Audio Processing*, 2004.

Chung-Hsien Wu and Yeou-Jiunn Chen, “Recovery of False Rejection Using Statistical Partial Pattern Trees for Sentence Verification,” to appear in *Speech Communication*, 2004.

Chung-Hsien Wu and Gwo-Lang Yan, “Acoustic Feature Analysis and Discriminative Modeling of Filled Pauses for Spontaneous Speech Recognition,” *Journal of VLSI Signal Processing*, Vol. 36, pp. 87-99, 2004.

Hunze Chen and Chung-Hsien Wu, “Speech enhancement based on audible noise spectrum and short-time spectral amplitude estimator,” *Electronics Letters*, Vol. 38(10), pp.485-486, 2002.

Jhing-Fa Wang, Jia-Ching Wang, An-Nan Suen, Chung-Hsien Wu, and Fan-Min Li, “VLSI Architecture and Implementation for Speech Recognizer Based on Discriminative Bayesian

- Neural Network,” *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E85-A(8), pp. 1861-1869, 2002.
- Yeou-Jiunn Chen, Chung-Hsien Wu, Yu-Hsien Chiu, and Hsiang-Chuan Liao, “Generation of Robust Phonetic Set and Decision Tree for Mandarin Using Chi-square Testing,” *Speech Communication*, Vol. 38(3-4), pp. 349-364, 2002.
- Chung-Hsien Wu and Jau-Hung Chen, “Automatic Generation of Synthesis Units and Prosodic Information for Chinese Concatenative Synthesis,” *Speech Communication*, Vol. 35, pp. 219-237, 2001.
- Chung-Hsien Wu and Jau-Hung Chen, “Automatic Generation of Synthesis Units and Prosodic Information for Chinese Concatenative Synthesis,” *Speech Communication*, Vol.35, 2001, pp.219-237.
- Chung-Hsien Wu and Yeou-Jiunn Chen, “Multi-Keyword Spotting of Telephone Speech Using a Fuzzy Search Algorithm and Keyword-Driven Two-Level CBSM,” *Speech Communication*, Vol. 33, pp. 197-212, 2001.
- Chung-Hsien Wu, Yeou-Jiunn Chen, and Gwo-Lang Yan, “Integration of Phonetic and Prosodic Information for Robust Utterance Verification,” *IEE Proceedings, Vision, Image and Signal Processing*, Vol. 147, pp. 55-61, 2000.
- Jen-Tzung Chien and Jain-Ray Lai, “Use of microphone array and model adaptation for hands-free speech acquisition and recognition,” *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, Vol. 36(2), pp. 141-151, 2004.
- Jen-Tzung Chien and Chih-Hsien Huang, “Bayesian learning of speech duration models,” *IEEE Transactions on Speech and Audio Processing*, Vol. 11(6), pp. 558-567, 2003.
- Jen-Tzung Chien, “Linear regression based Bayesian predictive classification for speech recognition,” *IEEE Transactions on Speech and Audio Processing*, Vol. 11(1), pp. 70-79, 2003.
- Jen-Tzung Chien, “Quasi-Bayes linear regression for sequential learning of hidden Markov models,” *IEEE Transactions on Speech and Audio Processing*, Vol. 10(5), pp. 268-278, 2002.
- Jen-Tzung Chien, “A Bayesian prediction approach to robust speech recognition and online environmental learning,” *Speech Communication*, Vol. 37(3-4), pp. 321-334, 2002.
- Jen-Tzung Chien, “Adaptive hierarchy of hidden Markov models for transformation-based adaptation,” *Speech Communication*, Vol. 36(3-4), pp. 291-304, 2002.
- Jen-Tzung Chien, Chih-Hsien Huang and Shun-Ju Chen, “Compact decision trees with cluster validity for speech recognition,” *Proceeding of International Conference on Acoustics, Speech, and Signal Processing*, Vol. 1, pp. 873-876, Orlando, May 2002.
- Jen-Tzung Chien, “On-line unsupervised learning of hidden Markov models for adaptive speech recognition,” *IEE Proceedings - Vision, Image and Signal Processing*, Vol. 148(5), pp. 315-324, 2001.
- Jen-Tzung Chien and Guo-Hong Liao, “Transformation-based Bayesian predictive classification using online prior evolution,” *IEEE Transactions on Speech and Audio Processing*, Vol. 369(4), pp. 399-410, 2001.

Jen-Tzung Chien and Ming-Shung Lin, "Frame-synchronous noise compensation for hands-free speech recognition in car environments," *IEE Proceedings - Vision, Image and Signal Processing*, Vol. 147(6), pp. 508-515, 2000.

Jen-Tzung Chien and Jean-Claude Junqua, "Unsupervised hierarchical adaptation using reliable selection of cluster-dependent parameters," *Speech Communication*, Vol. 30(4), pp. 235-253, 2000.

Jen-Tzung Chien, "Online hierarchical transformation of hidden Markov models for speech recognition," *IEEE Transactions on Speech and Audio Processing*, Vol. 7(6), pp. 656-667, 1999.

(三) 多模式人機介面

Ze-Jing Chuang and Chung-Hsien Wu, "Multi-Modal Emotion Recognition from Speech and Text," Accepted by International Journal of Computational Linguistics and Chinese Language Processing, 2004.

Ze-Jing Chuang and Chung-Hsien Wu, "Emotion Recognition from Textual Input Using an Emotional Semantic Network," *Proceedings of International Conference on Spoken Language Processing*, Denver, 2002.

Ze-Jing Chuang and Chung-Hsien Wu, "Emotion Recognition Via Acoustic Features And Semantic Contents In Speech," *Proceedings of International Symposium on Chinese Spoken Language Processing*, Taipei, 2002.

Ze-Jing Chuang and Chung-Hsien Wu, "Text-To-Visual Speech Synthesis For General Objects Using Parameter-Based Lip Models," *IEEE Pacific Rim Conference on Multimedia*, pp.589-597, 2002.

Ze-Jing Chuang and Chung-Hsien Wu, "Text-to-Visual Speech Synthesis for General Objects Using Parameter-Based Lip Models," *Lecture Notes in Computer Science Series, Springer-Verlag*, 2001.

(四) 自然語言處理

Jui-Feng Yeh, Chung-Hsien Wu, Ming-Jun Chen and Liang-Chih Yu, "Automated Alignment and Extraction of Bilingual Ontology for Cross-Language Domain-Specific Applications," *accepted by The 20th International Conference on Computational Linguistics (COLING'04)*

Jui-Feng Yeh, Ming-Jun Chen and Chung-Hsien Wu, "Semantic Inference based on Ontology for Medical FAQ Mining," *Proceeding of IEEE Natural Language Processing and Knowledge Engineering*, pp.710-715, 2003.

Chung-Hsien Wu and Gwo-Lang Yan, "Flexible Speech Act Identification of Spontaneous Speech with Disfluency," *Proceedings of European Conference on Speech Communication and Technology*, Switzerland, Geneva, 2003.

Yu-Sheng Lai, Kuao-Ann Fung and Chung-Hsien Wu, "FAQ Mining via List Detection," *Proceedings of the Workshop on Multilingual Summarization and Question Answering*, 2002. (A post-conference workshop in conjunction with COLING 2002)

Chung-Hsien Wu, Gwo-Lang Yan, and Chien-Liang Lin, "Speech act modeling in a spoken dialogue

system using a fuzzy fragment-class Markov model,” *Speech Communication*, Vol. 38, pp183-199, 2002.

(五) 人臉辨識

Jen-Tzung Chien and Chia-Chen Wu, “Discriminant waveletfaces and nearest feature decisions for face recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24(12), pp. 1644-1649, 2002.

Chih-Pin Liao, Hsien-Jen Lin, Chien-Yu Hung and Jen-Tzung Chien, “A new approach to face recognition with limited training data”, *Proc. of 2002 15th IPPR Conference on Computer Vision, Graphics and Image Processing (CVGIP)*, Hsinchu-Taiwan, pp. 46-50, August 2002. (in Chinese)

Jen-Tzung Chien and Chia-Chen Wu, “Discriminant waveletfaces for face recognition,” *Proc. of 2001 14th IPPR Conference on Computer Vision, Graphics and Image Processing Workshop (CVGIP)*, Pingtung-Taiwan, August 2001.

(六) 資訊檢索與分類

Yu-Sheng Lai and Chung-Hsien Wu, “Meaning term extraction and discriminative term selection in text categorization via unknown-word methodology,” *ACM Trans. on Asian Language Information Processing*, Vol. 1(1), pp.34-64, 2002.

Jen-Tzung Chien, Meng-Sung Wu and Hua-Jui Peng, “Latent Semantic Language Modeling and Smoothing,” *International Journal of Computational Linguistics and Chinese Language Processing*, August 2004.

Jen-Tzung Chien and Hung-Yi Chen, “Association rule based language models for discovering long distance dependency in Chinese,” *Research on Computational Linguistics Conference XIV (ROCLING XIV)*, pp. 43-63, Tainan-Taiwan, August 2001. (in Chinese)

自然語言處理與圖書資訊學

陳光華

國立台灣大學圖書資訊學系

khchen@ntu.edu.tw

學科領域的發展有其一定的理論基礎與脈絡，就圖書資訊學而言，其前身為圖書館學，依據美國圖書館學會的定義，其研究重點在於「將記錄下來的資訊加以選擇、蒐集、組織以及利用，以滿足特定使用者資訊需求的知識與技術。」(註 1)然而自 1960 年代起，電腦科技的普遍應用與資訊科學的興起，使圖書館學的研究典範逐漸擴展。1970 年代以來，圖書館學與資訊科學的整合更是熱門議題，致使 1980 年起，全球各國從事圖書館研究的機構，相繼掛上「圖書資訊學」的招牌，以揭示前述的轉變。不可諱言，圖書館學研究典範的擴展或變遷，實受到資訊科技與學科交流頻繁的影響。然而，若僅單純將其以單向傳播的角度來剖析這些轉變，亦有失之狹隘之憾。

長久以來，圖書館學、資訊科學、資訊工程學一直有非常緊密的關係。圖書館學著重於資訊載體 (Information Carriers) 與資訊之間的關係，同時，更注重如何提供良好的資訊服務 (無論是自動化服務或非自動化服務) 給予使用者 (讀者)。換言之，如何讓載體、資訊、與使用者三者間，產生良好互動，是圖書館學與圖書館員努力的目標，也是圖書館學研究的核心。至於，資訊科學則著重於資訊內在本質的研究，如齊普夫經驗法則 (Zipf Law)，夏儂之資訊理論 (Shannon's Information Theory)，以及所謂的「大世界悖理」引發的後續研究。至於資訊工程學則企圖以電腦做為工具，研究發展有效處理資訊的技術，希望藉由電腦執行，以取代人力或是降低人力的負擔。

為了讓研究工作得到更大的進展，學者專家無法侷促一隅，必須瞭解其他學科領域的理論與觀點。因此，即使不是全面的，上述三種學科領域也早就相互滲透與交流，只是 20 世紀末的發展加速了這項進程，使得有百年歷史的「圖書館學」這門學科領域的名稱，發生了根本的變動。

既然，圖書資訊學與資訊科學以及資訊工程學如此的緊密，其間的關係到底為何呢？本文旨在向資訊工程學的前輩、同儕、以及朋友們說明圖書資訊學與您們從事的研究的關係，因此，將略過圖書資訊學與資訊科學之間千絲萬縷的關係。此外，對於從事圖書資訊學研究的學者們，本文可說是班門弄斧，如有任何疏失，還請圖書資訊學界的前輩們見諒。

我想先說明個人對於圖書資訊學範疇的看法，當然圖書資訊學領域的學者專家有不同的見解，我認為圖書資訊學的內涵至少應包含以下幾點：

- 資訊徵集過濾
- 資訊處理技術
- 資訊系統管理
- 使用行為研究
- 分類編目索引
- 電腦網路技術
- 資訊仲介服務
- 資訊傳播研究

乍看之下，其中的資訊處理技術、電腦網路技術、與資訊系統管理，似乎應該是資訊工程學的範疇，然而圖書資訊學著重的與資訊工程學不同，如電腦網路技術中的通訊協定部分是圖書資訊學界關心的重點，包括 Z39.50 的訂定，OAI (Open Archive Initiative)，DOI (Digital Object Identifier) 等等。

不論圖書資訊學的範疇為何，圖書資訊學源生於圖書館學，在策略上，其與資訊工程學界有非常不同的切入點，圖書資訊學界重視的是資料、資訊、與知識；資訊工程學界重視的是自動化的處理技術，企望能夠研發 data-independent 的技術。此外，圖書資訊學有一個非常根深蒂固的取向，也就是「權威控制」的取向，然而這卻是資訊工程學界一直以來比較難以接受的論點。何謂權威控制？例如，一本圖書經由採購進入圖書館後，必須經由分類編目的過程，而後才能上架，讓讀者借閱。分類編目是一個權威控制的過程，圖書館員必須賦予該圖書一個分類號（如國會圖書分類法、中國圖書分類法、杜威圖書分類法之分類號），著者號，主標題目（權威控制的主題關鍵詞，用於描述圖書的內涵），以及其它的資料，也就是依據一定的 Metadata Format 給予 Metadata。而後使用者或讀者可以使用線上公用目錄系統（Online Public Access Catalogue，簡稱 OPAC）或是書名卡、著者卡、標題卡等紙本目錄來查詢所需圖書。換言之，需要耗費人力與時間來提供使用者「有用的」、「加值」後的資訊。

相對的，同樣的情況在資訊工程學界是如何處理的呢？假設該圖書為電子書，書的內容皆為數位資料，因此，可以使用資訊檢索的技術，例如利用 TF*IDF 的方式直接採用文件詞彙，計算詞彙的權重，然後以倒置檔的方式製作索引檔，提供使用者檢索。盡量避免人力的介入，完全以電腦自動化技術處理，快速簡便。表一簡單地比較前述二種作法。

表一 檢索系統的比較

| | 圖書資訊學界 | 資訊工程學界 |
|------|----------------|-------------|
| 系統內涵 | 目錄系統 | 內容系統 |
| 執行方式 | 人工製作 | 機器製作 |
| 詞彙選擇 | 控制詞彙 | 自然詞彙 |
| 詞彙組合 | 前組合（索引時組合） | 後組合（檢索時組合） |
| 優 勢 | High Precision | High Recall |

這個例子很清楚的說明「控制的觀點」與「自動的觀點」如何影響著圖書資訊學界與資訊工程學界的研究典範。

然而，令人感到欣慰的是，學術研究的發展越來越蓬勃，學術領域的交流越來越頻繁，跨越學科領域的合作研究越來越多，促使著學者專家更加理解其他學科領域的觀點、精神、與取向。對於圖書資訊學界的朋友而言，資訊已經多到無法都採用人工的方式處理，而且資訊的價值亦因時、因地、因人而有所不同，實不必使用相同的代價處理它；對於資訊工程學界的朋友而言，亦逐漸理解權威控制的價值，否則不會有 IEEE 與 ACM 聯合籌辦 Metadata 相關學術會議的舉動，也不會有 Semantic Web、Topic Map、與 Ontology 的研究。

對於我們計算語言學學會的朋友，以及學會秘書長交代給我的任務而言，計算語言學或是自然語言處理在圖書資訊學研究的角色才是本文的重點。（註 2）

基本上，計算語言學處理的語言可以區分為口語（Spoken Languages）與書面語（Written Languages），口語的研究離圖書資訊學比較遠，然而，書面語的研究與圖書資訊學的研究卻密切相關。圖書資訊學所要面臨的資訊載體，有紙本、類比、數位，若要採用自動的技術，如計算語言學的技術，必須將先紙本以及類比資訊轉為數位資訊。當然，有許多自動化或半自動化的方式可以進行這樣的轉換，同時也有許多自然語言處理的技術可供使用（如 OCR 時，採用語言模型增加辨識率），惟為便於論述說明，以下的討論皆假設我們面臨的皆是數位資訊。

前述論及圖書資訊學界與資訊工程學界的交流，當然會促使圖書資訊學界有新的觸發，就資訊檢索而言，採用統計為基礎的文件內容檢索，或是人工處理為基礎的文件目錄檢索，都不能滿足使用者資訊檢索的需求，因此，引入自然語言處理的技術，將文件內容視為書面語，進行深層的文件處理，提供更為精準的自動化資訊檢索系統，或是更有智慧的資訊服務

系統方是正途，畢竟圖書資訊學研究的最終目的是提供使用者最好的資訊服務，無論這樣的資訊服務是由何種方式達成的。

如前述，文件被視為書面語的載體，這種書面語的實質內涵即是吾人認知的「文本」。相對的，資訊檢索（Information Retrieval）的研究者不從語言角度入手，而將文件視為是「裝載詞彙的袋子」（Bag of Words），嘗試用統計的方法，擷取文字的特徵以建構資訊檢索的模式。因此，自然語言處理與計算語言學的技术可說是深層處理（Deep Processing）的技术，而資訊檢索的技术是淺層處理（Shallow Processing）的技术。

我們可以用以下的例子來說明自然語言處理對於資訊檢索的協助。若我們想要檢索下列新聞文件：「1997年6月28日，泰森在向霍利菲爾德 WBA 冠軍挑戰時，因不滿對方屢次擁抱和頭撞，而咬了對手的耳朵，...」。一般使用者會使用「泰森 耳朵」進行檢索，比較有耐心的使用者會使用「泰森 耳朵 霍利菲爾德」或是「泰森 咬 耳朵 霍利菲爾德」。這樣的檢索方式，會將前述新聞文件檢索出來，假設後來亦發生「霍利菲爾德也咬了泰森耳朵，作為報復」的新聞，甚或「霍利菲爾德在餐廳享用鵝肝醬燴豬耳朵，觀看泰森大戰強森的現場直播」也會被檢索出來。因此，採用 Bag of Words 的作法有其限制，若要提供精準的檢索服務，自然必須使用自然語言處理的技术。

對於某些圖書資訊中心，我們若要提供更高層次的資訊服務，如文件探勘，則自然語言處理的技术益形重要。因為一般文件的內容包羅萬象，全賴文件作者行文遣詞，無既定的規律，這類文件雖非「隱性知識」，但若欲以自動化的方法處理，卻也是相當大的挑戰。

文件有一定的結構，由上至下為：文件，段落（Paragraph），句子（Sentence），子句（Clause），詞組（Phrase），詞（Word），字（Character）。從語意的角度，段落之上，還有論域（Discourse），論域是指具有特定主題的口語或書面語，論域可能由一個以上的段落構成。文件結構的每一層次都有相應的處理程序：段落層次有段落的辨識與論域的處理；句子層次（與子句層次）有句子的辨識，有述語參數結構（Predicate-Argument Structure）的處理，並進一步建構其邏輯正規形式（Logical Normal Form）；詞組有詞組的辨識；詞有詞的辨識（Segmentation），詞類標記（Tagging），詞幹處理（Stemming），詞義標記（Sense Tagging）。這些程序僅是進行文件探勘的前置處理作業，真正要進行文件探勘時，還必須伴隨著分類或推論等技术，以探勘未知的知識。

無論是精準的資訊檢索服務或是智慧的文件探勘服務，會牽涉眾多的自然語言處理技术，以下簡要說明。

1. 文件結構的辨識

文件結構的辨識是處理非結構化資料的第一步，換言之，吾人必須給定文件每一成分所屬的結構標籤，如題名，作者，單位，機構，電子郵件，摘要，第一段落，其他段落，最後段落，圖片，表格等等結構標籤，至於到底有多少標籤，則由需求決定。文件的類型當然會影響文件的結構，因此，使用適應性的系統（Adaptive System）是較為可型的作法。

2. 文件論域的辨識

一篇文件可能有數個論域，一個論域可能有數個段落，辨識文件的段落後，必須辨識哪些段落是屬於同一個論域，這就是所謂的文件論域的辨識。論域具有核心的主題，非結構化資料的文件探勘或多或少都牽涉到語意層面的處理，因此，確認該核心主題是重要的工作。

3. 句子的辨識

句子的辨識是之後詞彙處理的基礎工作。句子辨識不如想像中的簡單，Palmer 與 Hearst 使用很複雜的處理程序以辨識英文句子，正確率為 98.5%。（Palmer & Hearst, 1994）中文句子的特性又與英文不同，中文使用短句很頻繁，通常以逗號「，」結尾。句子辨識完成後，需以句子為單位建立剖析樹（Parsing Tree），隨著語法理論的不同，剖析樹有眾多形式，

但是建構剖析樹之前，必須先處理詞組與詞彙。一旦建構完剖析樹，才能著手建構邏輯正規形式，而這是一切文件探勘中推論的基礎。

4. 詞組的辨識

詞組在英文中經常出現，通常一個相同動詞與不同的介副詞組合，就有不同的意義，這造成詞組的辨識是不可或缺的步驟。相對而言，中文詞組現象較少，倒是字組較多，如「一元復始」與「三陽開泰」。無論是詞組或是字組，其意義通常是個別的詞義或字義組合而兼有變化。目前使用多連（N-Gram），搭配統計模式以處理詞組或字組。如果有一部完整的片語詞典，亦可採用詞典為本（Dictionary-Based）的作法。

5. 詞彙的處理

中、日、韓等東方語系語言與西方的拉丁語系語言不同，東方語言有分詞（Word Segmentation）的問題，亦即東方語言詞彙間無西方語言詞彙間的空白標記。詞彙辨識完成，就是給定詞彙的詞類標記與詞義標記，另外，英文有詞幹處理的工作，中文則無。

6. 文件歸類

文件歸類與資料歸類相似，皆是希望將文件給定一個既定的分類標籤。這在新聞機構的應用特別明顯，一般的新聞機構都有其分類體系，如政治、財經、體育、娛樂、文學、旅遊等分類標籤。目前已經有許多的歸類辦法，只要建構待歸類文件的特徵，再據之給定最適當的分類標籤即可。

7. 文件分群

將文件依據各自的特性分群，而無既定的分類體系，通常分群的作法是採用階層式分群法（Hierarchical Clustering），最後所形成的群是由相似性的門檻值（Threshold）決定的，而決定門檻值並不容易，除了經驗與實作外，還沒有系統化的方式。分群的另一問題是如何給定分群的標籤或是建議分群的標籤，以揭示各分群的特性，這是文件分群最難的工作。

8. 自動推論

自動推論是人工智慧研究的一環，其本身就是很重要的課題，自動推論需要一套推論機制（Reasoning Mechanism），接受各文件擷取的資料，並作為推論的前提，運用推論引擎，以得到推論的結果。

9. 主題偵測與追蹤（Topic Detection and Tracking, TDT）

TDT 嘗試由一群依時間排序的文件，偵測新事件（新主題）的出現，並追蹤該事件，換言之，TDT 將文件依照不同的主題，將相同主題的文件依時序串連，建構時間序列，可探知事件的發源、演變、以迄終止。

10. 資訊擷取（Information Extraction）

資訊擷取是由文件中擷取事先預設所需的資訊；資訊檢索則是由文件集中檢索相關的文件。因此，「資訊檢索」這個詞彙事實上誤導了吾人對於相關研究的認識，如果「必也正名乎」，應該使用「文件檢索」代替「資訊檢索」。資訊擷取可視為比資訊檢索更深一層的資訊服務。資訊擷取不僅僅辨識重要的個體，還必須決定個體之間的關係。然而因為資訊擷取工作的特殊性，所以到底擷取何種資訊是依資訊服務系統服務的範疇而定。

11. 自動摘要（Summarization）

自動摘要則是以自動化的程序製作原始文件的精緻版。若從自動摘要模型的角度檢視所謂的自動摘要，可以分為兩種作法：第一種可由文件中挑選適當的段落或句子構成摘要，亦即製作所謂的「摘錄」；第二種則可由分析原始文件的角度出發，抽取文件的「概念表意」（Conceptual Representation），再進行「摘要的產生」（Summary Generation）。這兩種作法

各有其優缺點，基本上，第二種作法牽涉所謂「文件理解」的過程，若能夠真正達成所謂的「理解」，應該可以製作品質較高的摘要。然而對於網際網路的應用，時間是一個非常重要的限制條件，而理解通常必須花費相當長的時間。

必須再次強調圖書資訊學研究的最終目標是提供資訊使用者最好的服務，若參照前文提及圖書資訊學的範疇，您會發現整個研究範疇皆環繞此一核心目標。「資訊徵集過濾」是為了採訪並過濾適合使用者的資訊；「分類編目索引」是為了讓資訊各就各位，處於有序的狀態，讓使用者有效地使用資訊；「資訊處理技術」是針對不同的資訊需求，有效地處理各式資料，以提供妥適的資訊服務；「電腦網路技術」是為了整個資訊服務系統的基礎建設；「資訊系統管理」是讓系統可以有效以及有效率地運轉；「資訊仲介服務」是作為使用者與資訊的中介，服務無法獨立取得所需資訊的使用者；「使用行為研究」可以讓資訊服務的提供者，更加瞭解使用者的行為，以設計更好的資訊服務系統；「資訊傳播研究」則是探討資訊傳播的模式，可作為資訊服務的參考。我們可以發現其中的「資訊徵集過濾」、「分類編目索引」、「資訊處理技術」與自然語言處理密切相關，前述的相關技術皆可以運用於圖書資訊學研究，以提供更佳的資訊服務。

在目前學科領域交流越來越頻繁的情況下，作為一位研究人員，若能瞭解所處學科領域的大框架，以及與相關學科領域的連結，就比較能夠理解各學科領域的底層結構與基本取向，從而做出正確的判斷，整合各學科領域研究的優勢，進行跨越學科領域的合作研究。

註釋

註 1：Young, Heartsill ed. (1983). The ALA Glossary of Library and Information Science. Chicago: American Library Association, p.132.

註 2：對於許多嚴謹的學者與專家而言，自然語言處理與計算語言學有很大的不同，本文並不特意強調其間的異同，而會交替使用這二個術語。