網路論壇 FAQ 知識之自動轉換設計

陶幼慧
高雄大學資訊管理系

黃清俊
精業公司

楊誌欽
高雄海洋科技大學微電子工程系

摘要

從現代知識管理的角度而言，常見問答集(Frequently Asked Questions，FAQ)可謂一種將原始討論文章內容經過篩選與整理，並轉換成可分享的特定表達格式的知識內涵。網路討論區大量的資訊，需要密集的人力投入，方能彙整成符合讀者需求品質的FAQ。事實上，斷詞與自動摘要等相關技術之研究與應用已相當成熟與普及，網路討論區的 FAQ 知識應該可自動轉換產生，讓 FAQ 變成討論區的基本功能之一。基於上述動機，本研究提出一網路討論區 FAQ 知識自動轉換的概念模式，而為探討該模式之可行性，並實作了一個 FAQ 自動摘要的雛型系統以進行實證。第一階段實驗評估顯示，在經過篩選與整理的五個討論區主題，壓縮比為 20%時其摘要之平均召回率與精確率高於 45%，與相關文獻中之表現相當；而 FAQ 摘要的提示性、可讀性、句數適當性及介面的輔助性的評估上，獲得很高的認同度。於第二階段實驗未經篩選的五個討論區主題時，雖然多篇回應文章的討論主題發散，使得摘要的精確率與召回率降低，但互動式介面的便利，讓使用者主觀的接受度仍高。因此，本研究兩階段實驗結果，展現了本研究所提之 FAQ 知識轉換的概念模式，在技術整合與使用者觀點上的初步可行性。然而，於討論區回應文章發散的特性上，仍需進一步地整合自動摘要相關技術，以改善摘要文章的結構與提示完整性。本文最後並討論自動摘要實務應用上的意涵，以及未來繼續研究的方向，以期 FAQ 知識轉換的概念模式可以成功地被應用於知識管理的企業組織中。

關鍵字：常見問答集、多文件摘要、資訊檢索、知識管理、自動摘要。

Designing an Automatic FAQ Abstraction for Internet Forum

Yu-Hui Tao
Department of Information Management, National University of Kaohsiung

Ching-Chun Huang
Systex Corporation

Chih-Chin Yang
Department of Microelectronic Engineering, National Kaohsiung Marine University

Abstract

From the perspective of modern knowledge management, frequently asked questions (FAQ) is a knowledge that is screened and organized from original articles, and then transformed into a standard question-and-answer representation for easier sharing. An Internet forum contains a large volume of interactive information, which requires extensive manpower input to maintain its readability and quality. In fact, related research and applications in text retrieval or automatic abstracting have been quite matured and popular, such that automatic FAQ abstraction can be a basic function of any Internet forum software. Accordingly, we proposed a conceptual model for automating the FAQ generation, and a prototype system was implemented to explore the feasibility of this model. In the first phase, five human-adjusted Internet forums were experimented, and the recall and precision were both over 45% under 20% of compression rate, which were equivalent to the averages seen in the literature. Moreover, the evaluation on indicative of content, readability, appropriateness of sentence size, and interface assistance showed a very high acceptance. In the second phase, five unadjusted Internet forums were experimented, the recall and precision rates were low due to the divergent character of Internet articles, but the user acceptance was still high due to the friendly interface. Overall speaking, the two-phase experiment supported the technical feasibility and user acceptance of the proposed model, but the divergence of Internet articles needs to be dealt further. Some implications and future work are also discussed at the end.

Keywords: Frequently Asked Questions (FAQ), Multiple-Article Abstracting, Information Retrieval, Knowledge Management, Automatic Abstracting

壹、緒論

網際網路工具中以討論區、電子佈告欄或新聞論壇(NetNews)類等非同步互動方式，最易凝聚人氣而形成網路社群，故其在全球資訊網上的應用更為一般企業網站所偏好，而其應用也逐漸普及於教育上與一般生活中。長久以來，常見問答集(Frequently Asked Questions，FAQ)已經成為這些網路社群既有之參考資訊來源，而 FAQ 往往是由社群管理者或熱心成員負責彙整討論內容重點，並以問答方式摘錄並集結成文字檔，提供其他成員或新加入者的一個便利方式，以取得該社群相關討論主題之重點資訊。相對地，文字資訊也帶來其他的問題，無用且大量的資訊會造成了「資訊需求者」的負擔，如討論區、留言版經常出現一些無意義的聊天文章，有過量或品質參差不齊的資訊，並未帶來相對的好處。大部分的電子佈告欄、留言版與討論區，都只是作為聊天或沒有固定主題的討論，因此容易產生過多無意義的文章；「主題討論區」因係針對有興趣的議題進行討論，甚至於其中之「網路論壇」更往往限制在某一特定期間內討論，所以在內容方面，相較於一般無主題電子佈告欄或討論區而言，不適當的內容也相較減少，但討論文章的主題發散的問題並未完全消失。大量的文字資訊同樣地亦造成了「資訊管理者」的負擔，要整理如此龐大的資料，需要花費非常多的人力與時間，隨著討論區等類似工具的普及，已氾濫到很高比例的管理者疏於管理或甚至於不去管理的現象，反而造成 FAQ 知識形式的相形式微。在知識管理意識抬頭的大環境中，企業甚至於個人都體會到知識分享的好處與價值，這樣一個既有的 FAQ 知識轉換與分享現況，更需要關注與保存。

知識管理的定義多樣化，綜合一般文獻的觀點，知識管理的精神在於擷取、儲存、轉換及傳播組織內的知識(Carliner 1999; Teece 1998; Holsapple & Joshi 1999;譚大純等人 1999)。網路社群本身符合知識之擷取、儲存與傳播的程序，而 FAQ 的製作彌補了網路社群在知識轉換及分享程序上的不足。因此，注重知識經濟的現代企業，在實施知識管理時往往喜歡以文件管理系統管理顯性知識，或以討論區等社群網站形式營造隱性知識分享的文化與環境。

由上述討論得知，網路討論區已經成為企業知識管理應用的一個主要工具，但長久使用於網路社群中的 FAQ 卻因過度普及的副作用而逐漸式微。主題討論區的文章與 FAQ 的製作實乃文件摘要的轉換工作，現今的文件自動摘要技術相關研究已有相當的應用水準。因此，如何快速地整合現有資訊技術來輔助網路社群管理者，自動產生 FAQ 知識，是一個值得關心的議題。本研究針對網路討論區 FAQ 自動產生的研究問題，首先進行相關關鍵技術的文獻探討，並從實務應用之導向，建構一個自動摘要 FAQ 的概念模式，為驗證該模式之技術可行性與使用者的接受程度，並將快速開發一雛型系統進行一實驗性的評估。

貳、文獻探討

本研究文獻以國內、外相關之文件自動摘要技術為主，進行完整的文件自動摘要的概略探討後，並就其中與本研究相關的中文斷詞、句子相似度、與績效衡量的部份做進一步的探討。

一、文件自動摘要

文件摘要發展的範疇是從傳統單篇文章摘要到多篇文章摘要(沈建誠 2001)或多國文字摘要(蘇哲君 2001)。文件自動摘要因其目的所需，一般有提示性(indicative)、訊息性(informative)、重要性(critical)及比較性(comparative) 四種格式(Ando et al. 2001)，而文獻中最多的應用為訊息性摘要，強調原文中重點的完整性，其次為提示性摘要，提供讀者決定是否閱讀全文的參考決策。文件摘要過程一般包括分析、選擇、濃縮及表達四個主要步驟(Marybury 1995)。其中「分析」包含了字頻 (word frequency)、線索片語 (clue phrases)、佈置 (layout)、語法、語意、語段 (discourse)與語言使用(pragmatics)；「選擇」包含了字頻、線索片語、統計與結構 (structure)；「濃縮」包含了抽象化 (abstraction)及加總 (aggregation)；「表達」包含了計畫、實行 (realization)及佈置。

Zhi et al. (2001)指出自 1958 年第一個自動文章摘要演算法出現以來，相關的方法主要源自兩個應用領域，即以統計方法導向的資訊檢索(Information Retrieval, IR)及以自然語言處理為基礎的資訊摘錄(Information Extraction, IE)。IR 之下又分為傳統及語料庫(corpus-based)兩種途徑，其中傳統方法以字頻等表面特徵的評分規則確認重要(salient)訊息，而語料庫則增加了複雜的語料庫訓練機制。IE 之下也分為語段及知識豐富(knowledge rich)兩種途徑，其中語段以一些文字內容的了解取代傳統個別表面特徵的方式，知識豐富則利用應用領域的已知知識去轉化摘要結構及組成。文章摘要內容篩選原則上以句子為單位，往上提昇至內文段落(Salton et al. 1997; 黃純敏 2001)，進而至多文段落或句子，因此單文摘要與多文摘要類似(翁鴻加 2001)，只不過多文摘要在方法運用上因範圍較廣而有較多的步驟處理程序(黃思萱 2002；沈建誠 2001)。

自動摘要可用於如新聞摘要、報章產品評論摘要及技術文件摘要等情境(Zhi et al. 2001)。近十年的發展已成熟至開發雛型應用系統並驗證其技術效益之層次，例如 Lam & Ho (2001)的智慧型財務新聞摘要系統(Financial Information Digest System, FIDS)、Zhi et al.(2001)開發以代理人為基礎的文字摘要系統 iTSum 作為解析財務與股市新聞為例、Lehmam (1999)的以敘述片段為基礎的自動摘要(Automatic Summary based on Indicative Fragments, RAFI)以科學及技術文件為摘要對象、Brandow et al. (1995)的自動新聞摘要系統(Automatic News Extraction System, ANES)以特定的 41 種發行刊物為摘要對象，及吳潮崇(2002)自動摘要線上新聞，按照新聞事件的性質予以分類，以同一類的新聞事件製作成一份摘要。一般而言，目前較易成功的自動摘要應用為有特定應用領域對象及其相關知識輔助的情境，這也說明了通用性摘要技術的方法無論多麼強大，其績效也依賴自製或市場現有之語料庫。

Brandow et al. (1995) 指出要自動產生可讀性高的摘要，必須深度了解文章內容、確定資料間之相關重要程度及產生關聯性的輸出結果。一些文獻研究中也都有訂定其摘要考量，例如：沈建誠 (2001)從多篇文章摘錄一篇摘要的角度，以滿足指示性簡單摘要，和查詢主題相關兩條件為主，而能因應使用者的查詢而有所改變；邱立豐(2002)

從資訊檢索往往忽略關鍵字間之關係，評估「相關詞組」應用在資訊檢索時，對檢索者在關鍵詞擴展的效益。Gong and Liu (2001)認為好的一般性摘要應該包含文章主要的主題並將重複多餘降至最低；Neff and Cooper (1999)的主動摘要與標記系統(Active Summarization and Markup, ASHRAM)強調能自動選擇、標記及連結有用的資訊，讓使用者在瀏覽時能不需閱讀所有文件即可較輕易地抓到重點；Lam & Ho (2001) 開發 FIDS 允許使用者進行交叉確認其原始全文內容；Ando et al. (2001)利用關鍵字萃取文章摘要，以輔助使用者決定是否閱讀本文；Knight & Marcu (2002)探討以機率方法於句子壓縮，以滿足文法正確及包含重要訊息兩個假設。

目前商業用途的摘要往往摘錄文章最前面的句子作為摘要，例如 Lycos 搜尋引擎提供前幾百字元作為該文章之摘要(Neff & Cooper 1999)，Mead Data Central (MDC)的 Searchable Lead 資料庫也提供前 60 字、150 字及 250 字為該文章之摘要(Brandow et al. 1995)。從文獻研究發現，最佳文章摘要比例約為 20% (Morris et al. 1992)，Neff & Cooper (1999) 提及自己實驗及 Verity 軟體開發者都指出至少要四句話才讓使用者能夠主觀地接受這樣的摘要。Goldstain et al. (1999) 的研究也建議一篇文章的摘要至少要三至五句話。

在技術限制下，基本方法應用之外，一些研究也強調系統整體實用性。如 Lehmam (1999) 的 RAFI 強調其簡易使用、友善介面及快速、並能滿足使用者需求。Neff & Cooper (1999) 的 ASHRAM 強調提供一個主動輔助使用者瀏覽文件的前端介面設計。所以，其介面搜尋的結果除了有原始全文外，並提供關鍵字、摘要及連結到本文的超連結。Rothkegel (1995)提出一個三層式文字模式，分內容、功能到格式剖析文章，該模式較適用於結構性較強的領域，例如該文中引用科學領域的文件來說明內容、功能與格式的規則。

二、中文斷詞

自動化文件摘要的首要步驟就是斷詞，而中文文件斷詞主要有統計式斷詞(Sproat & Shih 1990)、詞庫式斷詞(Chen & Kiu 1992)及混合式斷詞(Nie, Briscbois & Ren 1996)三種。統計式斷詞，需藉由大量的文件或語料庫的統計資料(詞頻、門檻值)，以鄰近字元出現的頻率高低來決定斷詞的位置。優點是不需專家定義詞彙，執行效率高，能有效解決詞庫在擷取複合名詞、專有名詞及新生詞彙的問題；但如果詞長大於二時，則斷詞的效率會大幅降低，以及提高演算法的複雜度(Nie, Briscbois & Ren 1996)。由於語料庫與應用領域有關，不同語料庫間的統計資料不適合互用，極易影響斷詞的正確性。

詞庫式斷詞，是目前最為普遍使用的斷詞方式，主要是利用現有詞庫與文件進行比對，一般最常使用的方式為長詞優先法。然而詞庫式斷詞的缺點受到詞庫品質的影響很大，對於複合詞彙、大多數的專有名詞與新生詞彙無法辨識，為了要提高斷詞的正確性，必須不斷的更新詞庫的內容。混合式斷詞，將詞庫式斷詞法與統計式斷詞法整合(Nie, Briscbois & Ren 1996)，先利用詞庫式斷詞找出不同的斷詞組合，再利用詞

彙的統計資料，找出最佳的斷詞組合。另外，陳稼興、謝佳倫與許芳誠(2000)提出以遺傳演算法(Genetic Algorithms)為基礎的中文斷詞模型，用以處理中文斷詞的問題。其透過文章的訓練自行產生詞庫，其優點可避免人為介入的不客觀性，也可避免浪費人力資源。

在資訊檢索技術中，詞彙的擷取是一門重要的技術，相關的應用領域有胡勝傑與許中川(1999)、黃燕萍與許中川(1999)及陳景揆與許中川(2001)等運用混合式斷詞針對新聞文件擷取每篇新聞文件的關鍵詞彙。黃純敏(2001)及黃純敏與吳郁瑩(1999)等以詞庫式斷詞作為關鍵字萃取工具，擷取網路上的超文件自動產生摘要。黃聖傑(1999)以統計式斷詞進行多文件自動摘要的研究。翁頌舜、杜宜潔(2002)利用詞庫式斷詞擷取新聞關鍵字，來建置個人化的新聞資訊推薦系統。黃國豪與歐仁德(2002)以遺傳演算法建置詞庫，製作郵件自動回覆客服系統。由上述的文獻可以知道中文斷詞技術於各個應用領域都有相當的研究。

三、句子相似度計算

文件摘要的目的是將重要的句子摘錄成一篇簡短的文章，輔助使用者能夠在較短的時間內了解文章的內容，為了要將重要的句子結合成摘要，需要評估句子的重要性。Edmundson (1964) 提出了關鍵詞法、位置法、線索法及標題法四種方法作為句子重要性評估準則。黃純敏與吳郁瑩(1999) 搭配位置法、標題法及網路標籤判斷，來評估句子的重要性，但其權重的設定涉及太多的主觀認定及不確定因素，所以一般研究還是以較為客觀的關鍵詞法作為評估句子的重要性的準則(邱立豐 2002)。

一般評估文章的重要性可考慮詞彙在文章中所出現的頻率及關鍵詞彙所在的位置，以推估其文章的重要性。在計算句子的權重方式主要分為 TF*IDF(Term Frequency /Inverse Document Frequency)及相似度兩種計算方式，TF*IDF 的計算方法，其字詞的權重與詞頻成正比，而其出現的文章篇數成反比，依此方法計算出的字詞的權重值作為評估句子重要性的依據。句子的相似度計算，則是利用句子間關鍵詞重疊多寡來求句之間的相似度大小，一般會設定一個門檻值作為判斷句子之間是否達到高相似度值，而在文章中與較多句子具有高相似度時，相對地代表此句子在文章中愈重要。

由上述可知，相似度的大小是藉由兩句子皆存在之關鍵字所佔的比重大小而決定，在計算相似度的方法主要有 Dice、Jaccard 及 Cosine 三種係數(Brandow et al.1995; Singhal et al. 1996)。沈健誠(2001)於多篇文件摘要系統對句子的評估，主要考慮兩方面，包含提示性與主題相關性(topic relevance)。提示性的計算是以兩句子間相同的動詞與名詞的重複次數，而句子的提示性又分為內部提示性與外部提示性，一個句子的評分包含了內部提示性、外部提示性與查詢主題相關性。Salton et al. (1997) 將文字關係分布圖 (text relationship map) 的概念應用在文件摘要的研究上，其從文件連結觀點探討百科全書內文段落之關聯性，並提出以「段落」為摘錄單位的文件摘要系統，希望藉由相關段落的組合以克服文句連貫性的困難。

Salton 在相似度的組成摘要上，提出了 Global Bushy Path、Depth-First Path 及 Segmented Bushy Path 三種內文連結模式。黃純敏(2001)將三種內文連結模式，經評估實驗結果，發現其中 Global Bushy Path 在鏈結的一致性及內容廣度都比另外兩個方法有較好的成效。於是採用 Global Bushy Path 的方法，並以句子為擷取的比較對象，依據關鍵詞彙共同出現的頻率，引用 Jaccard coefficient 計算句子間的相似度值。全篇文章的分句權值，則為該句與其它句的相似度總和，亦即連結點越多的句子權值愈高，代表該句越重要。

四、績效評估

文件摘要的評量指標以召回率(recall)、精確率(precision)、壓縮比(compression)及接受度(acceptability)為主。召回率將人工評估認為重要的句子(H)，作為文章相關句子總數的基數，系統所擷取到與人工擷取相同的重要句子(C)作為分子，計算系統選句的召回率(C／H)；精確率則將系統所擷取到的重要句子(M)作為分母，對照系統所擷取正確之相關句子總數的基數(C／M)。文章最佳摘要並不具唯一性，不同的目的或不同的人進行摘要都可能導致不同結果。客觀的評估以自動摘要的結果與人工摘要結果做一比較，忽略多重正確答案的可能性；主觀的方式則評估摘要的接受程度 (Neff and Cooper 1999)。

文獻中以精確率與召回率評估的範例最多：Gong & Liu (2001)的通用性摘要方法評估採用召回率、精確率與兩者之平均，其平均值介於 52%與 61%之間；Brandow et al. (1995) 的 ANES 系統評估的精確度約為 45%，而召回率約為 56%；Ziu et al. (2001) 的 iTSum 系統評估結果整體而言，於壓縮率 25%下及三句摘要兩種情形下，精確率約為 50%左右與召回率約為 70%左右；Ando et al. (2001) 的日文複合字摘要的精確度介於 45% 與 63%間；Marybury (1995)的 SumGen 系統評估因以資料庫事件為摘要來源，故其精確率與召回率高達 90%以上，其實驗設計並包含任務操作時間等其它衡量指標；Chien (1997)其召回率與精確率約為 43%與 30%；葉鎮源(2000)兩種摘錄方法，於壓縮比 30%下，召回率分別為 52%及 45.6%。

Brandow et al. (1995) 的 ANES 系統評估主觀之接受度介於 68%-78%之間，並與 MDC 商業產品 Searchable Lead 的結果作比較，整體而言，ANES 的接受度雖不低，但卻遠遠低於 MDC 的 Lead 平均 92%的結果。令人深思的一點卻是 Lead 的摘要僅是取文章前面的 60 字、150 字或 250 字而已，很明顯的僅具提示性，不保證訊息完整性，但卻能贏得極高的接受度。沈建誠 (2001) 則從多篇文章摘錄一篇摘要的角度，以滿足指示性簡單摘要，和查詢主題相關兩條件為主，其實驗所得到的摘要縮減比率為 95％以上。

參、研究方法

本研究問題是在探討如何將網路論壇上的文章，透過自動化的流程，彙整並摘錄文章內容重點，進而產生 FAQ 的知識形式，以提供網站管理者或討論區版主輕鬆地及有效率地進行知識擷取與轉換的工作環境，進而實現知識分享的目標。因此，研究步驟將先依文獻整理結果彙整出一多文件 FAQ 摘要之概念模式。並實際開發一雛型系統，進行概念模式之實驗評估。從上述自動摘要相關文獻探討得知，自動摘要在較簡單的非語意技術方法的應用，已相當成熟且應用廣泛。因此，本研究並非以技術研發導向，而為技術整合之實務導向，以期讓自動產生討論區 FAQ 的知識管理應用，具實質上的管理意涵與導入之基礎。

就目前文獻探討了解所作的假設，包含將針對內容較為嚴謹的「主題討論區」多文章摘要為範疇，製作通用性摘要；摘要格式與目標為兼具提示性與訊息性，亦即各篇文章有意義的內容必須條列，所有篩選句子之訊息完整性以最大為考量；強調實務可行性成果，對於查詢摘要內容的需求層面，提供原文之超連結與關鍵字等訊息追蹤介面。然而，本研究最大的限制在於，目前網路上一般的論壇，雖有特定主題，但多為不限時間的開放式討論，造成單一主題的多篇回應文章，不但文章結構鬆散，而且內容易與原文的主題偏離，形成發散的現象。因此，與現有文獻中偏向單文摘要，或文章來源為結構嚴謹的網路新聞或技術文件相較，預期本研究在客觀評估的精確率及召回率將不及文獻中看到的平均績效評估結果。因此，使用者的主觀接受程度，是實務導向的本技術整合研究的重要指標。

以下分別介紹本研究提出的 FAQ 知識轉換的概念模式、雛型系統設計、以及兩階段實驗設計。

一、FAQ 知識轉換的概念模式

本節分別說明 FAQ 知識轉換概念模式中的高層次運作機制、摘要通用流程以及兩階段相似度計算之架構。針對主題討論區建構的運作機制如圖 1 所示，主要分主題討論區、斷詞比對模組及 FAQ 摘要三個模組。

主題討論區包含討論區內容及提供給使用者的 FAQ 摘要；斷詞比對模組運用詞庫及語料庫進行討論區文章內容之斷詞比對，以產生關鍵詞資訊，另外在詞庫式斷詞之外，本機制亦納入統計式斷詞以產生詞庫中不存在之新生詞資訊；FAQ 摘要模組則擷取關鍵字詞與新生詞資訊，經由文章句子的相似度計算組織 FAQ 摘要的內容，並彙整 FAQ 摘要以適當之表達互動方式呈現給使用者資訊。

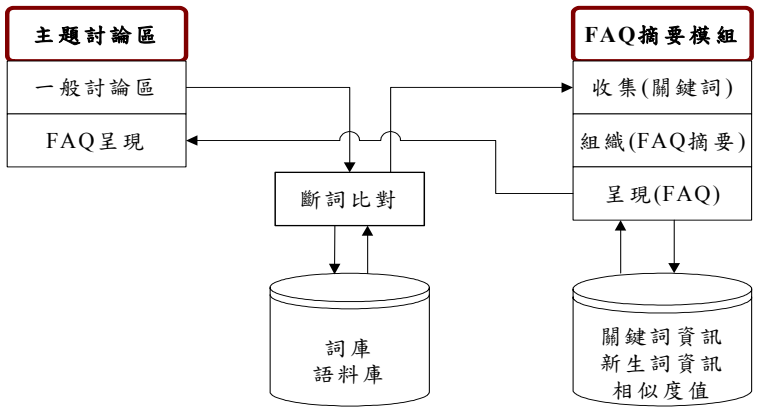

圖 1:主題討論區 FAQ 摘要運作機制 (資料來源：本研究)

基於高層次的運作機制，本研究彙整之多文件 FAQ 摘要流程著重其一般之通用性，因此將含括一些技術方法於相關過程表達之中，以期於不同主題討論區的需求，可自訂選擇適用之技術方法。如圖 2 所示，多文件 FAQ 摘要完整之通用流程可分為輸入、斷句、斷詞、單文件句子相似度計算、單文件重要句子篩選、多文件句子相似度計算、多文件重要句子篩選及 FAQ 產生等八個步驟。相對於文件摘要過程的四部分：分析包括斷句與斷詞；選擇與壓縮包括單文件句子相似度計算、單文件重要句子篩選、多文件句子相似度計算及多文件重要句子篩選；表達為 FAQ 產生。

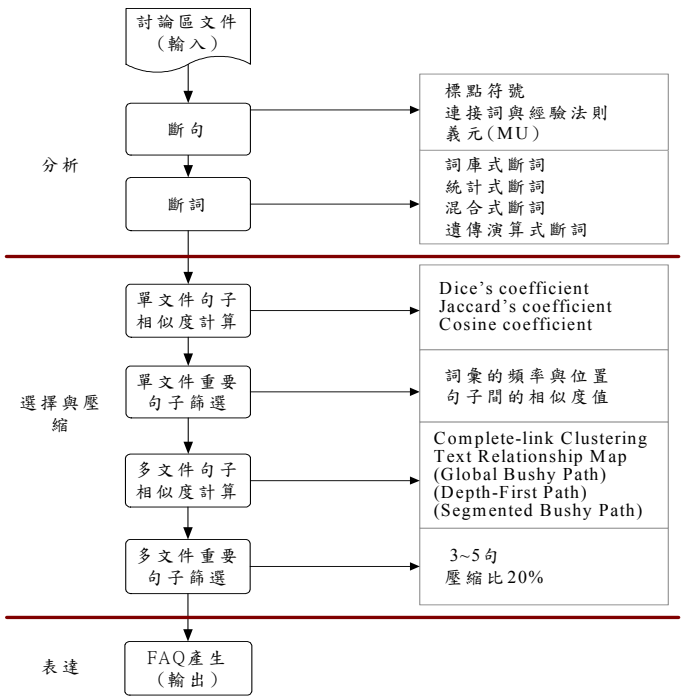

圖 2:多文件 FAQ 摘要通用流程 (資料來源：本研究)

FAQ 摘要通用流程中相似度計算部分，因本研究為多文件摘要性質，將相似度計算分兩階段進行，如圖 3 所示。兩階段計算過程中，首先進行各單篇文章相似度的計算，求得各篇文章中各個句子的相似度值；在第二階段中，取得各篇相似度值較高的句子，進行多篇文章的句子相似度值計算，在多篇文章相似度值較高的句子，彙集成為 FAQ 的摘要結果。兩階段相似度比較的優點在於當文章數目比較多的時候，每篇文章所篩選出的句子數較少，在第二階段跨文章句子評估時效率比較快；相對的，所有文章的所有句子進行一次的相似度評估，雖然沒有句子在評估過程中被遺漏或是先淘汰，但執行效率將隨著文章句字數成長而快速降低。因此，本研究採取兩階段相似度評估的考量，主要在於網路討論區文章成長快速時，「即時執行效率」是使用者滿意接受度的主要考量，稍微降低的自動摘要內容的精確性是可以接受的。

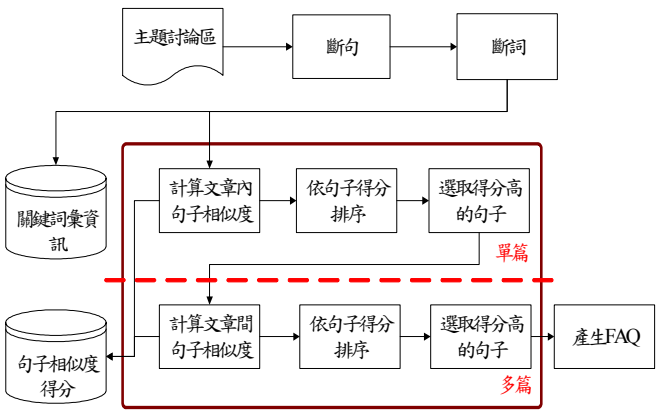

圖 3：兩階段相似度計算（資料來源：本研究）

二、雛型系統

雛型系統的實作考量，於斷句方面，以句號(。)、分號(；)、驚嘆號(！)、問號(？)及逗號(，)作為分隔符號，並需同時考慮到半形與全形兩種情況。在斷詞方面，本研究僅將動詞與名詞視為與文章內容最相關的重要詞彙(黃純敏、吳郁瑩 1999)，採取詞庫式斷詞與統計式斷詞擷取新生詞彙的混合式斷詞法，其中詞庫比對結合長詞優先法進行斷詞作業，而既有詞彙的擷取以 2 字詞到 9 字詞為主，並以長詞為優先選取對象，比對中央研究院之八萬詞目詞庫中收集的詞目。擷取新生詞彙使用統計式斷詞的基本概念，經由詞庫式斷詞後比對不出的未知詞，新生詞彙即在這些字詞的集合中，透過統計式斷詞找出未知詞彙的組合，再依其出現的頻率決定是否為新生詞彙，達到擷取新生詞的目的。

本研究採取黃純敏(2001)所提依 Salton 的內文連結觀念，採用 Global Bushy Path 的方法，以句子為實驗的對象，並以 Jaccard coefficient 法則計算句子間的相似度值，

依據句子中關鍵詞彙共同出現的頻率，作為衡量句子間相似度，並且採取兩階段式的相似度計算。計算公式如下：

$$A\text{句與}B\text{句的相似度} = \frac{Z}{X+Y-Z} \qquad \text{----- (公式 1)}$$

其中，X 表示 A 句中關鍵詞數目，Y 表示 B 句中關鍵詞數目，Z 表示 A、B 句中重複的關鍵詞數目。依照 Jaccard coefficient 在求得各句之間的相似度值後，相似度值會介於 0 至 1 之間，全篇文章的分句權值，則為該句與其他句的相似度總和，亦即連結點越多的句子權值越高，代表此句子越重要。為提昇使用性及降低摘要不完整下之不滿意程度，摘要的表達，將以重要句子條列之，並可連結閱讀原文，並且將關鍵字條列於摘要之前，讓使用者可有充分瀏覽關鍵字、摘要及全文之自由度及彈性。

本研究所提出之 FAQ 摘要雛型系統建置於微軟 Windows 2000 Server 作業系統與 IIS 5.0 網站伺服器，並搭配 ASP 網路程式語言開發。資料庫部分因使用中央研究院之八萬詞目詞庫，經簡單處理後儲存在 Microsoft SQL Server 2000 的資料庫中，供斷詞比對使用。

本研究雛型系統主要畫面如圖 4 所示，以 A、B 與 C 三塊解說。上方 A 區為論壇主題「如何在平凡的生活中發現藝術」，主題下方的關鍵字「藝術、生活、環境」則為相關文章中最重要的語詞中篩選而出。左方 B 區的多文摘要內容，則為各文章中篩選出的重要句子，句子之間的空格為文章的分野。右方的 C 區顯示 B 區各句的說明，包含第幾篇文章的第幾句、及該句子中的關鍵字。

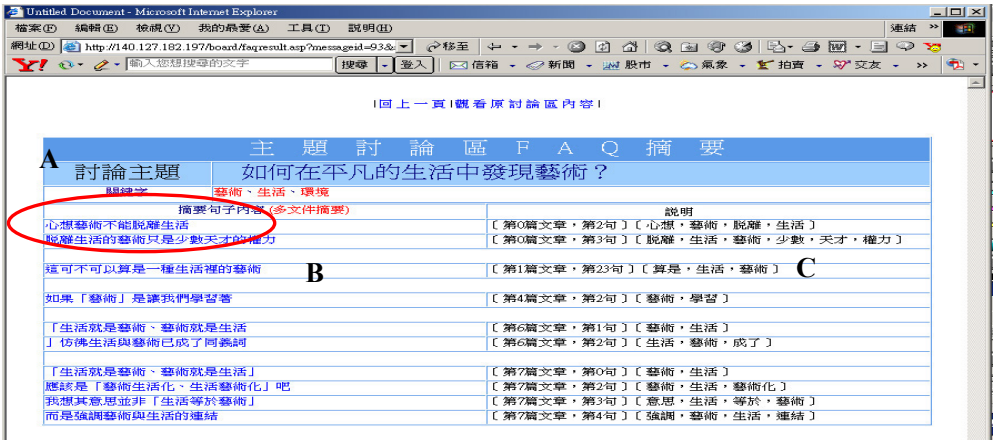

圖 4：多文件 FAQ 摘要雛型系統畫面 (資料來源：本研究)

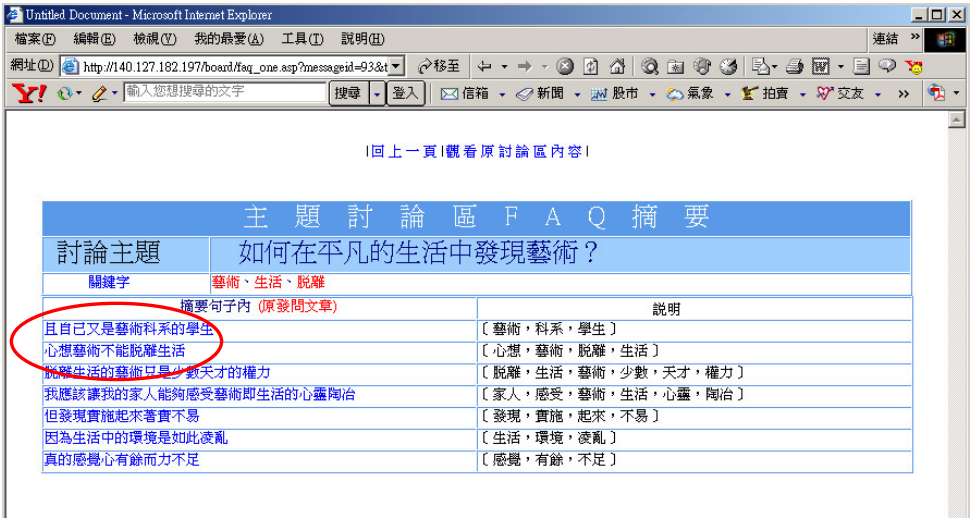

圖 5：單文件 FAQ 摘要畫面（資料來源：本研究）

點選 B 區的句子，將可超連結至該文章的單篇摘要，指標並將停留在該句上。單篇摘要的格式與圖 4 的多篇摘要格式類似，惟點選單篇摘要的句子時將超連結至該篇原文，供讀者仔細閱讀該文章細節，例如圖 4 中第一篇文章兩句之單文摘要如圖 5 所示，而該文之原文透過超連結的連結如圖 6 所示，連結的句子並以紅色標示，以便利讀者參照其相對位置。

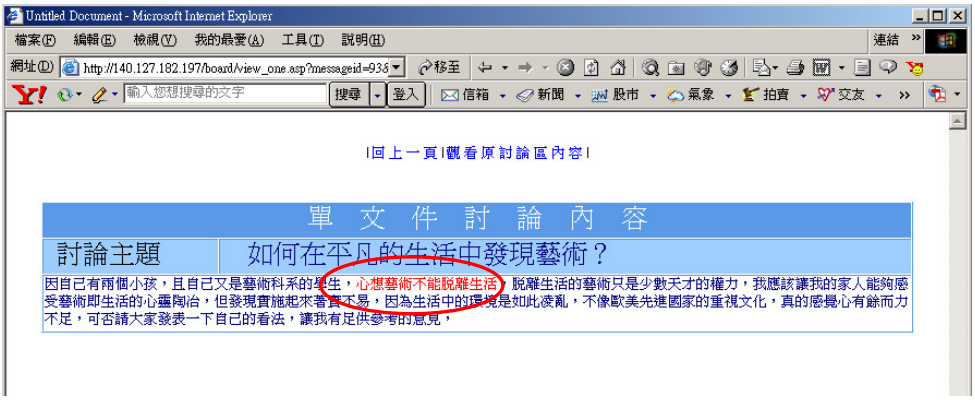

圖 6：單篇文章原文畫面（資料來源：本研究）

本系統程式開發過程，程式的驗證是透過單篇文章到多篇文章的資料量，採由下而上 (bottom-up) 的方式逐段的測試討論區文章的讀入、斷句的判斷、與八萬詞庫比對、相似度比較、到單篇與多篇摘要產生的過程。每階段的資料皆寫入程式中的陣列變數，部分重要過程資料並寫入資料庫中，供程式驗證所用。而前段所述之使用者介面，提供使用者從多文摘要中，查看關鍵字、摘要句連結回去單篇摘要與其關鍵字、以及單篇摘要連結回單篇原文，亦輔助了由上而下(top-down)的程式驗證的第二管道。綜合本系統開發的經驗，討論區自動摘要的概念模式與實作的設計考量，對系統程式開發人員是困難的，但在已提供自動摘要的功能規格情況下，整體的程式撰寫與

測試，從程式量與困難度來講，對熟悉軟體系統開發的專業人士而言，則是相當容易的工作。因此，本研究概念模式的建立與雛型系統開發的過程，在 FAQ 自動摘要的技術整合應用上，有其實務上參考的重要性。

三、第一階段實驗設計

本研究的兩階段實驗，均透過 FAQ 摘要雛型系統的操作說明，以及實際讓使用者進行文章的評選與系統使用，來收集使用者對文章所評選摘要與系統自動摘要所得結果，以分析人工評選重要句子與系統之間有無差異，驗證本研究所提之 FAQ 摘要系統的可行性。同時，搜集每位使用者對於 FAQ 摘要系統之主觀評估，進行分析探討。

第一階段的實驗目的，是希望在控制主題討論區文章的發散程度下，實證如果論壇的回應文章與原始文章的主題較為貼切時，其精確率與召回率應可維持與文獻探討中相當的水準。為達成這樣的目標，在方法的選擇，以簡易了解與開發，並具相當之績效為原則；雛型系統的介面應提供便利的人機互動與功能為原則；客觀評估之精確率與召回率期望在 45%以上，人為主觀判斷之接受度期望在 75%以上。

在實驗的文件數量上，收集 5 個一般性的討論主題的文章來進行實驗，每個討論的主題有 6 至 8 篇的文章，每個主題的句子數約為 100 句，如表 1 所示。討論主題是刻意篩選文章結構不會太鬆散，且回應文章也不致過於偏離原始主題，少數部分偏離句子也被刪除。

表 1：主題討論區的資料來源 (資料來源：本研究)

討論主題	網站	文章篇數	文章句數
1.善用工具	易學網---學習討論區	6	107
2.多點創意思考少點傳統束縛	易學網---學習討論區	7	109
3.如何自我訓練口語能力和寫作技巧	敦煌書局：外文書籍的專業網站【討論區】	7	106
4.創意思考	全球藝術教育網---藝教討論區	8	106
5.如何在平凡的生活中發現藝術	全球藝術教育網---藝教討論區	8	109

在實驗的人員數量上，設計為每個討論區主題由 3 人評估，共 15 位評估人員進行 FAQ 自動摘要效益評估的實驗，每位參與實驗的研究對象針對一個討論主題的文章進行評選。評估者首先必須仔細閱讀指定的討論主題，針對所閱讀的文章評選出具有代表性並與討論主題相關的句子。由於每個討論的主題由 3 個人進行評估，同一個討論的主題，被二人以上同時認為重要的句子數，將作為精確率與召回率計算的基準。在評選完畢時，請評估者瀏覽系統產生的 FAQ 摘要的結果，並採用傳統之問卷方式，要求評估者就系統自動摘要在壓縮比 10%、20%及 30%的提示性、可讀性、句數適當性及介面的輔助性等項目填答。問卷題目的尺度區間設計為 1 分~5 分(表示從非常不同意至非常同意)，以利後續簡易之計量分析與討論。

第二階段實驗設計

為了進行田野式主題討論區文章的實驗評估，本研究針對第一階段討論區的全球藝術教育網、易學網學以及敦煌書局：外文書籍的專業網站，與第二階段的思摩特網、Java 技術論壇以及藍色小舖六個討論區網站進行文章篇數抽樣調查與比較。兩階段的討論區資料來源與抽樣數如表 2 及表 3 所示，兩階段至少都抽樣 500 個以上主題以及 4500 篇以上文章。這六個討論區網站，都是相較於一般如雅虎或 MSN 入口網站，討論主題較為嚴謹且文章提供者都是較為認真地參與討論。

表 2：第一階段討論區網站抽樣資料來源 (資料來源：本研究)

討論區	主題數	文章數
1. 全球藝術教育網---藝教討論區 http://gnae.ntptc.edu.tw/arted/forum_list.jsp	237	7179
2. 易學網學---討論區 http://stu.esharer.com.tw/dis_group/list_all.asp	150	843
3. 敦煌書局：外文書籍的專業網站---英語教學主題討論區 http://www.cavesbooks.com.tw/Forum/List.asp?Forum=英語教學	191	1832

表 3：第二階段討論區網站抽樣資料來源 (資料來源：本研究)

討論區	主題數	文章數
1. 思摩特網---資訊科技融入學習領域教學討論區 http://sctnet.edu.tw/nineyear/SIGboard.php?dbn=Nineyear_cai	249	1138
2. Java 技術論壇--- Software Engineering 討論區 http://www.javaworld.com.tw/jute/post/page?bid=33&sty=1&age=0&tpg=5	120	1284
3. 藍色小舖---資訊安全討論區 http://www.blueshop.com.tw/board/list.asp?fumcde=FUM20041006152 9473A1	509	2535

文章篇數在抽樣的主題中所佔的比率分布如圖 7 與圖 8 所示。圖 7 大致分為 0-9 篇、10-20 篇、以及 21 篇之後三種斜率的趨勢，超過 60%的討論區主題的回應文章篇數落在 0-9 篇的第一種斜率類別中，也說明了第一階段實驗設計具有相當的代表性；圖 8 與圖 7 的三種斜率趨勢雷同。因此，第二階段實驗，延續第一階段的實驗設計基礎，選擇思摩特網與藍色小舖討論區中回應文章超過 30 篇的五個主題，如表 4 所示；並將每個主題依三種斜率的趨勢分為低(9 篇)、中(18 篇)、與高(25 篇)量三個族群，不作文章的任何處理，進行與第一階段相同之實驗程序。五個主題與三種討論主題文章數量，共有十五種組合；每種組合需要三位評估者，因此共需四十五位評估者。為便利高量族群進行關鍵字之選擇，本階段實驗並將討論區主題與回應文章列印出來，可直接讓評估者比較勾選關鍵字。

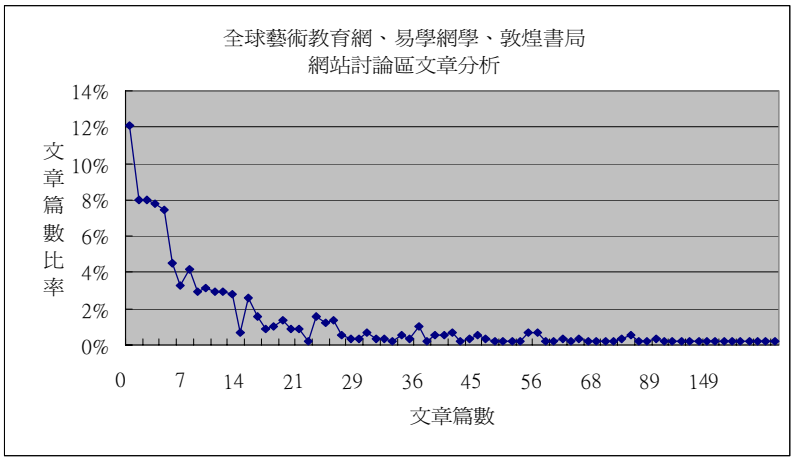

圖 7：第一階段實驗討論區網站分析（資料來源：本研究）

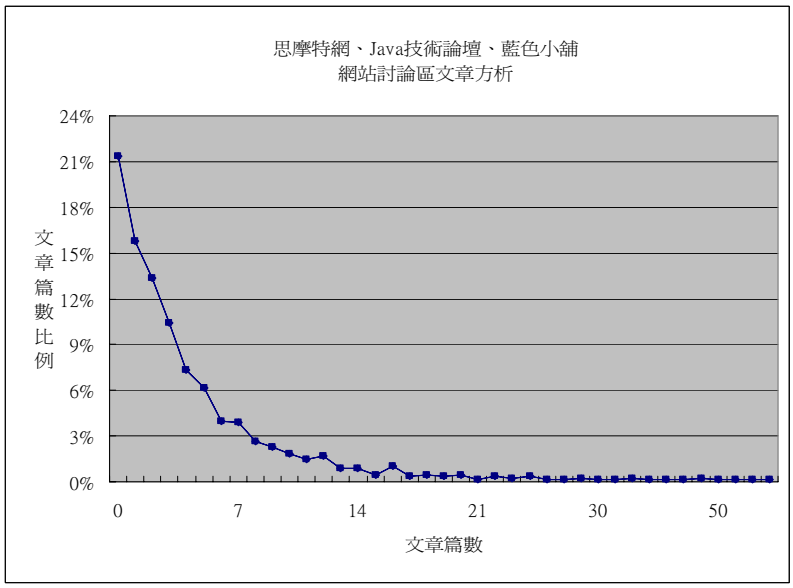

圖 8：第二階段實驗討論區網站分析（資料來源：本研究）

表 4：主題討論區的資料來源（資料來源：本研究）

討論主題	網站
1. 九年一貫是注定要失敗的	思摩特網---九年一貫的精神與內涵討論區
2. 什麼是美	全球藝術教育網---藝教討論區
3. 資訊科技融入教學的困境？	思摩特網---資訊科技融入學習領域教學討論區
4. 處罰算不算是體罰？	思摩特網---人權教育討論區
5. 何謂人文？	全球藝術教育網---藝教討論區

肆、第一階段實驗結果與分析

實驗評估的結果與分析，從客觀文字檢索衡量與主觀的使用者意見兩方向進行。因為第一階段的實驗目標在於證明，與討論區原始文章主題貼切的回應文章所自動產生的摘要，在客觀與主觀的評估指標下，至少都有文獻中相關研究的水準，所以結果的呈現與分析上均較為詳盡，以作為第二階段實驗比較的基礎。

一、不同壓縮比之精確率與召回率

實驗的結果顯示，在人工擷取重要句子數方面，人工擷取的重要句數約在 12 至 19 句，而系統擷取的重要句子數，依其壓縮比的不同各為 10 句、20 句與 30 句，與人工擷取的句數有一定的差距。如表 5 所示，依其壓縮比的不同，在壓縮比為 30%時，其召回率大多高於 50%，各為 83.3%、57.9%、53.9%、61.1%，僅主題 3 的召回率為 46.2%，但其精確率較為低，各為 33.3%、36.7%、20%、30%、36.7%；在壓縮比為 20%時，其召回率各為 75%、47.4%、38.5%、35.3%、55.6%，其精確率各為 45%、45%、25%、30%、50%；在壓縮比為 10%時，其召回率，各為 41.6%、26.3%、30.8%、23.5%、38.9%，其精確率大多高於 40%，各為 50%、50%、40%、40%、70%。

表 5：第一階段實驗之「召回率」與「精確率」 (資料來源：本研究)

	召回率			精確率		
壓縮比 / 主題	10%	20%	30%	10%	20%	30%
1	41.6%	75%	83.3%	50%	45%	33.3%
2	26.3%	47.4%	57.9%	50%	45%	36.7%
3	30.8%	38.5%	46.2%	40%	25%	20%
4	32.5%	35.3%	52.9%	40%	30%	30%
5	38.9%	55.6%	61.1%	70%	50%	36.7%

以上實驗結果的解釋為當系統擷取的句數越多，其精確度則逐漸降低，故其精確度最高僅 70%，最低達到 20%；在召回率的部分，當系統擷取的句數越多，系統擷取的句子與人工擷取的句子的重疊機率越高，所得到重疊的句數就會越多，因此壓縮比為 30%時的召回率會高於壓縮比為 20%與 10%時，其召回率最高為 83.3%，最低為 23.5%。，不同壓縮比下的召回率與精確率的差距，部分反應了 Neff & Cooper (1999) 所謂的忽略多重正確答案的可能性。

依過去文獻研究的結果，摘要最佳的壓縮比率約 20%左右；而依其摘要的性質與摘要內容的不同，其召回率與精確率約在 40%至 60%左右。而本研究結果發現，在壓縮比為 10%時，精確率大多高於 40%但其召回率較低。在壓縮比為 20%時，主題 1、2 與 5 的召回率與精確率皆高於 45%，僅主題 2 與 4 之精確率雖不高，其召回率也接近 40%左右。在壓縮比為 30%時，其召回率大多高於 50%，但由於系統擷取的句數較高，所以導致其精確度較低。因此，在召回率以 30%最佳，精確率以 10%最佳，然而 10%與 30%的召回率與精確率相斥。從平衡的角度來看，20%的召回率與精確率的結果較適中，且與文獻的數據相當。

二、多文件 FAQ 自動摘要效益評估項目

本研究之多文件 FAQ 摘要的目標乃兼具提示性與訊息性，其目的在於輔助使用者判斷主題討論區討論之原文大意。若要閱讀單篇的摘要、單篇原文或討論的全文，使用者可進一步的利用本系統提供之「超鏈結」功能。為了輔助使用者能更清楚地了解討論的內容，介面中亦呈現摘要與句子的關鍵字等資訊。因此，效益評估的項目分為兩部分，主要評估使用者對系統 FAQ 摘要結果，以及系統所提供介面輔助的接受度。

(一)摘要結果評估

FAQ 摘要結果之評估項目，包估摘要的提示性、可讀性及句數適當性等三個問項，在不同的壓縮比之下，各項目的接受度的結果如表 6 所示，衡量尺度為 1 分的非常不同意到 5 分的非常同意。其中「提示性」與「可讀性」的項目，在壓縮比為 30%時的平均數為最高，其平均數為 4.33 與 3.93，然而 20%其各為 4 及 3.8，與 30%之差異不大；而「句數適當性」是在壓縮比為 20%為最高，其平均數為 3.93，然而與壓縮比 30%之 3.87 差異也不大。

表 6：FAQ 摘要之評估 (資料來源：本研究)

評估項目	問題	壓縮比		
		10%	20%	30%
提示性	FAQ 摘要能有效表達原文大意	3.67*	4	4.33
可讀性	FAQ 摘要所摘錄之語句連貫	3.47	3.8	3.93
句數適當性	FAQ 摘要所摘錄之長度適當	3.53	3.93	3.87

*李克氏量表 5 尺度：1 表示「非常不同意」到 5 表示「非常同意」

以上實驗結果的解釋為當摘要壓縮比越大，其摘要內容也越多，FAQ 也就提供較多之資訊，所以使用者的同意傾向會越高，使得在壓縮比 30%時的「提示性」與「可讀性」之評比最高。反過來看「句數適當性」，摘要內容越多即代表摘要的長度增加，理論上會造成句數的適當性降低結果，然而，可能是本階段實驗討論區文章內容總長度僅在 100 句左右而已，故壓縮比 30%的同意程度雖然降低，但 10%也不是最佳的句數，反倒是壓縮比 20%時的句數最被受試者接受。

總而言之，本實驗的壓縮比在 30%與 20%的「提示性」、「可讀性」及「句數適當性」評比結果相當且均可接受，而平均值為李克氏量表 5 尺度的 4 左右，具正面高接受度的意義。如果與召回率即精確率的結果共同考量時，20%的壓縮比會是一個較佳的選擇。

(二)系統介面之輔助性

系統介面之輔助性，主要是評估本實驗之雛形系統所提供的關鍵字與超鏈結等功能，是否能讓使用者在瀏覽 FAQ 摘要時，更加容易的了解文章的內容？各問項的接受度的結果如表 7 所示。其中「超鏈結的功能」之平均數為 4.13，「關鍵字的提示」之平均數為 3.13，「語句之關鍵字」之平均數為 3.53，「介面格式」之平均數為 4.07；在此評估項目中「超鏈結的功能」之平均數為最高，而「關鍵字的提示」之平均數較低。

表 7：系統介面之輔助性評估 (資料來源：本研究)

問題	平均數*
超鏈結的功能能夠有效的輔助了解討論區的內容	4.13
關鍵字的提示能夠有效的輔助了解討論區的內容	3.13
摘要語句之關鍵字等資訊能夠有效的輔助了解討論區的內容	3.53
摘要之介面格式能清楚的閱讀摘要的內容	4.07

*李克氏量表 5 尺度：1 表示非常不同意到 5 表示非常同意

就整體而言，系統提供的輔助性與系統介面的接受度，有很高的滿意程度，其中「超鏈結的功能」與「介面格式」的平均數大於 4，不但具高度正面接受的意義，也反映出受試者對摘要結果的閱讀較重視，而對於摘要如何產生的線索，如全文之「關鍵字」及摘要語句中之「關鍵字」則相對重視較低。值得後續 FAQ 自動摘要於操作介面改善之重點方向參考。

伍、第二階段實驗結果與分析

為奠定兩階段實驗比較的共同基礎，第二階段實驗的結果呈現與分析比照第一階段實驗。綜而言之，第二階段召回率與精確率遠不如第一階段，然而使用者主觀的摘要評估與系統介面之輔助性評估仍保持相當高的水準。表 8 顯示客觀的 15 組召回率與精確率數據，基本上，不但大多遠低於第一階段的結果，且精確率更是偏低。更甚者，看不出來具體的趨勢或樣式可供判斷；僅能推論，影響主題討論區的主要因素就是回應文章與原始文章主題貼切的程度，至於回應文章的數量並無絕對的影響，顯示了多重正確答案的可能性(Neff & Cooper 1999)。在討論主題發散的情形下，影響技術性指標績效的程度更大。然而，召回率與精確率在壓縮比方面仍約是相斥的結果，因此採平衡角度的 20%壓縮比的建議仍可適用。

如表 9 所示，45 位評估者 FAQ 摘要的主觀評估，平均值在壓縮比 30%時的提示性、句數適當性、以及可讀性項目中，都是最高的。然而，壓縮比 20%時的平均值也都大於 3 且接近壓縮比 30%的水準。與第一階段實驗不同的是，10%的壓縮比的三個項目的平均值皆低於 3，象徵 10%的壓縮比在現有主題討論區的現況中並不適用。所以，20%或 30%的壓縮比都是可以被評估者接受的範圍。

表 10 顯示評估者對介面輔助性評估的四個主觀項目，平均值皆大於 3，與第一階段實驗相同的是，仍以「超鏈結」與「摘要的介面格式」為較高。關鍵字雖然有幫助，但使用者閱讀摘要時並不一定需要知道關鍵字。就使用者主觀判斷的最高兩項指標與第一階段結果相同，本研究可推論，儘管客觀的召回率與精確率並不理想，但現有的主題討論區在友善與瀏覽的互動式自動摘要環境下，仍可達到實務上使用者接受的程度。

表 8：第二階段「召回率」與「精確率」 (資料來源：本研究)

主題 \ 壓縮比		召回率			精確率		
		10%	20%	30%	10%	20%	30%
1	少量	20	60	63.2	14.8	14.6	13
	中量	36.2	73.2	80.6	18	10	11.6
	多量	23.2	31.6	70	32.8	31.4	23
2	少量	17.6	23.5	33.5	33.2	22.2	14.1
	中量	10.8	29.7	35.1	28.6	25.8	20.4
	多量	17.3	39.1	39.1	10	5.6	7.4
3	少量	26.6	33.3	46.6	24.2	19.2	18.4
	中量	7	14.2	16.5	16.8	9.3	8.5
	多量	14.8	18.5	29.6	23.2	16.8	17.8
4	少量	12.2	27.2	45.4	29	24.4	22.2
	中量	26.3	29	40.5	20.6	15.8	12.1
	多量	36.2	37.2	49	27.2	23.6	20.3
5	少量	20	42.3	44.2	23.2	20	17.6
	中量	8.3	43.2	52.6	25.2	21.3	7.4
	多量	9.4	47.6	66.6	23.4	28.8	16.6

表 9：FAQ 摘要之評估 (資料來源：本研究)

評估項目	問題	10%	20%	30%
提示性	FAQ 摘要能有效表達原文大意	2.77	3.33	3.82
句數適當性	FAQ 摘要所摘錄之長度適當	2.95	3.35	3.41
可讀性	FAQ 摘要所摘錄之語句連貫	2.28	3.2	3.48

*李克氏量表 5 尺度：1 表示「非常不同意」到 5 表示「非常同意」

表 10：系統介面之輔助性評估 (資料來源：本研究)

問題	平均數
超鏈結的功能能夠有效的輔助了解討論區的內容	3.62
關鍵字的提示能夠有效的輔助了解討論區的內容	3.26
摘要語句之關鍵字等資訊能夠有效的輔助了解討論區的內容	3.46
摘要之介面格式能清楚的閱讀摘要的內容	3.66

*李克氏量表 5 尺度：1 表示非常不同意到 5 表示非常同意

陸、結論與建議

本節先就自動摘要技術整合應用之研究結果作一簡單結論，然後針對管理上如何應用本研究成果作建議說明，以及列舉未來可能的研究方向。

一、結論

客觀指標評量的結果，於第一階段中討論區主題發散程度控制的情況下，壓縮比20%與 30%的召回率與精確率均達到文獻中相關研究的 45%平均水準，而從平衡召回率與精確率兩者的角度而言，20%壓縮比的結果是可以接受。在沒有控制討論區文章的第二階段實驗結果，召回率與精確率均不理想，且無法觀察出較顯著的趨勢。但從平衡召回率與精確率的角度，20%的建議仍是有效的。主觀衡量指標的部分，雖然第二階段實驗的結果受到客觀指標不佳的影響，但整體而言，兩階段實驗的結果一致：在接受度方面，評估摘要的提示性、可讀性、句數適當性及介面的輔助性等項目，在壓縮比為 20%與 30%時的接受度較佳且相當，與相關研究之文獻比較仍是可接受的；在介面的輔助性的評估方面，「超鏈結的功能」與「介面格式」之接受度平均值較高，因此整體的輔助性與系統介面的接受度，有很高的使用者滿意程度

二、管理意涵與建議

根據上述實驗結論，企業在實務應用上應有下列之正確認知，方能實現討論區 FAQ 自動摘要的預期效益：

1. **可直接適用在討論區文章較不發散的環境**：從本研究兩階段實驗成果可知，當討論區文章與原始文章主題貼切時，客觀的召回率與精確率與其它研究文獻中的水準相當；但無法控制文章發散程度時，客觀的衡量指標就很不理想。因此，企業可就現有主題討論區，做快速的盤點檢視，挑選一些長期以來，原始與回應文章間貼切程度在某個標準之上的主題，直接套用 FAQ 自動摘要的功能，立即可看到成效。
2. **迫切需要分享的討論區主題，也可嘗試互動瀏覽之環境**：從兩階段實驗結果可知，討論回應文章發散程度是好是壞，對於摘要內容與瀏覽摘要的互動介面兩個層面而言，使用者主觀認知上接受程度皆很高；在控制發散得宜的第一階段中，一半項目的平均值更超過 4.0。因此，企業組織內部一些影響層面較廣，且有迫切需要分享的討論區主題，應用 FAQ 自動產生的摘要互動式環境，應仍可獲得內部員工正面的肯定。例如，公司的專案小組，需要尋找公司內部其它專案的經驗時，或相關專業社群的討論內容時，FAQ 的取得可有效的提供判斷原始討論主題文章的適當程度，並快速做成深入閱讀與否的決定。所以，企業在進行討論區文章盤點檢視時，一些有助於其它員工或群體參考的討論區主題，又不很在意召

回率與精確率的技術指標時，便可套用本研究的 FAQ 自動摘要的功能與互動式的介面環境。

3. **掌握討論區 FAQ 自動摘要的特性，可開創一些新的應用情境**：企業如果了解本研究模式與實驗的結果，且又能掌握其特性的話，可開發一些適用的情境與活動，以創新企業內部知識分享的方法或形式。例如，在企業教育訓練的課程設計中，可指派特定主題，讓受訓學員於特定時間內針對原始文章以及已發布的回應文章，回應個人的看法，並於指定時間結束後，自動產生 FAQ 摘要。此種情境，不但參與學員的回應文章可獲得有效的凝聚收斂，在結訓後也可獲得精簡的討論摘要，更可提供後續受訓學員有效的參考。這可能是一般企業進行教育訓練未曾思考的方向，但有了這項工具，創新應用的情境即可產生。類似的創新應用情境，可應用於企業內部內聚力較強的專案組織、跨功能部分的工作流程、單一功能部門或者利益與興趣相關的群體中。

三、未來研究

本研究從結論與管理應用上，可看出四點未來研究方向：

1. **專業領域語詞庫的建立**：本研究雛型系統採用中研院開發的中文八萬詞庫，作為斷詞以及文章句子相似度判斷的依據。然而，中文八萬詞庫涵蓋的是一般通用的中文詞語，專業領域的詞語則較缺乏，故目前僅適用非專業領域的討論區主題。再反過來看，企業內部原本就會有許多專業領域討論主題，而且其文章收斂程度往往也較一般主題高，更容易是企業內部分享的標的主題。所以，本 FAQ 自動摘要的概念模式，有必要增加一專業語詞庫產生的功能，以擴大現有模式適用的範圍。專業語詞庫的模組，可應用如文獻探討中提及的統計式斷詞(Sproat & Shih 1990)或遺傳演算法(陳稼興等人 2000)等技術。
2. **發散主題融入摘要的改良**：從兩階段實驗中得知，影響召回率與精確率最重要的因素，是文章發散的程度。雖然，文章發散是討論區文章的一個特色，僅能從參與討論成員的配合度改善，但是發散的主題，卻可透過分群演算法，使摘要內容可涵括這些延伸的副主題，使摘要更具提示完整性。加上現有摘要結構，雖然使用者給予整體不錯的評比，然而如圖 4 所示，目前以文章及其句數出現之先後順序顯示，而且重要的句子容易重複出現，摘要的整體順暢度仍待加強。後續可參考 Edmundson (1964)的位置法、線索法及標題法，以及 Salton et al. (1997)內文段落關聯性，做進一步自動摘要呈現格式的改善，使 FAQ 自動摘要內容更像一篇獨立的整合性文章，並且涵蓋多篇回應文章延伸出的副主題。
3. **斷詞與自動摘要重要議題的考量**：第二階段低落的召回率與精確率技術指標，說明了這是一個困難但重要的研究議題。另外，本研究並未考慮正、負語意之方向，並無法精確判定文章內容與關鍵字正、負語意之適當程度；網路上濃縮、簡化、口語化等傾向，在語意擷取與判斷也是重要的議題之一。因此，長遠的後續研究目標，當針對這些斷詞與自動摘要的典型議題，更深入的探討融合整合現有的技術，以解決這些議題，並期望提昇更紮實的客觀績效指標基礎。
4. **個人化參數設定之應用**：本研究提議的概念模式已具相當之通用性，在相同的程式邏輯運作下，有相當的彈性空間可進行程式邏輯的部分參數控制，以提供更個人化

的使用者互動介面。主要的參數有摘要長度以壓縮比或句數的選擇、壓縮比例或句數的設定、相似度以兩段式或一次評估、給特定領域討論區使用之專有名詞語庫之建立、同義字語庫的建立與比對結果之合併等等。雖然部分參數超越本研究雛型系統與實驗的範圍，但個人化參數設定在技術可行性上並不困難，且讓使用者擁有自動摘要執行的部分主導權，反而更容易獲得較高的使用者接受滿意程度。

致謝

本研究承蒙國科會專題計畫 NSC 92-2416-H-214-002 之經費與人力補助，特此致謝。另外，也感謝國立屏東科技大學資管碩士班畢業的林志龍，協助本研究實驗的進行。

參考文獻

1. 沈健誠，2001，多篇文件自動摘要系統，國立清華大學資訊工程研究所碩士論文。
2. 邱立豐，2002，互動式概念查詢應用於網路文件自動摘要之效益，國立雲林科技大學資訊管理研究所碩士論文。
3. 吳潮崇，2002，線上新聞之自動摘要系統，國立台北科技大學電機工程研究所碩士論文。
4. 胡勝傑、許中川，1999，『中文新聞文件斷詞』，第十屆國際資訊管理學術研討會，國立中央警察大學，第 968-974 頁。
5. 翁鴻加，2001，多文件摘要一些新技術及評估模型之建立，國立台灣大學資訊工程研究所碩士論文。
6. 翁頌舜、杜宜潔，2002，『建構線上個人化新聞資訊推薦系統之研究』，第八屆國際資訊管理研究暨實務研討會，國立高雄第一科技大學，第 943-950 頁。
7. 陳稼興、謝佳倫、許芳誠，2000，『以遺傳演算法為基礎的中文斷詞研究』，資訊管理研究，第二卷，第二期，27-44 頁。
8. 陳景揆、許中川，2001，『探勘中文新聞文件』，資訊管理學報，第七卷第二期，103-122 頁。
9. 黃純敏，2001，多語文（中英文）超文件自動摘要與評估，行政院國家科學委員會專題研究計畫成果報告。計劃編號：NSC89-2416-H-224-053。
10. 黃純敏、吳郁瑩，1999，『網路文件自動摘要』，台灣區網際網路研討會 TANET'99，國立中山大學承辦。
11. 黃思萱，2002，以關鍵詞分群為基礎的多文件摘要，資訊管理研究所碩士論文，國立台灣科技大學。
12. 黃國豪、歐仁德，2002，『以遺傳演算法及模糊理論為基礎之多問題自動回覆客服系統』，第八屆國際資訊管理研究暨實務研討會，國立高雄第一科技大學，第 345-351 頁。

13. 黃燕萍、許中川，1999，『中文社會新聞文件資訊擷取』，第十屆國際資訊理學術研討會，國立中央警察大學，第 975-982 頁。
14. 黃聖傑，1999，多文件自動摘要方法研究，台灣大學資訊工程研究所碩士論文。
15. 葉鎮源，2000，文件自動化摘要方法之研究及其在中文文件的應用，國立交通大學資訊科學研究所碩士論文。
16. 譚大純、劉廷揚、蔡明洲，1999，『知識管理文獻之回顧與分類』，中華民國科技管理研討會，中山大學，623-635 頁。
17. 蘇哲君，2001，中英雙語多文件自動摘要系統研究，國立台灣大學資訊工程研究所碩士論文。
18. Ando, K., Qamasaki, T, Shishibori, M. and Aoe, J.-I., "Automatic text summarization based on keyword derivation", *IEEE International Conference on Systems, Man and Cybernetics*, 10/7-10/10, 2001, pp: 464-469, Tucson, AI, USA.
19. Brandow, R., Mitze, K. and Rau, L. F., "Automatic condensation of electronic publications by sentence selection", *Information Processing & Management* (31:5) 1995, pp: 675-685.
20. Carliner, S.," Knowledge management, intellectual capital and technical communication", *International Professional Communication Conference*, 1999, pp: 85-91.
21. Chen, K. J. and Kiu, S. H., "Word identification for mandarin Chinese sentences", *Fifth International Conference on Computational Linguistics*, Nantes, France, 1992, pp: 101-107.
22. Chien, L.-F., "PAT-tree-based keyword extraction for Chinese information retrieval", *Proceedings of the 20th Annual ACM SIGIR Conf. on Research and Development in Information Retrieval* 1997, pp: 50-58, Philadelphia, USA.
23. Edmundson, H. P., "Problems in automatic abstracting", *Communication of the ACM* (7: 4) 1964, pp: 259-263.
24. Goldstain, J., Kantrowitz, M, Mittal, V. and Carbonell, J., "Summarizing text documents: sentence selection and evaluation metrics", in *Proc. Of ACM SIGIR'99*, August 1999, Berkeley, CA.
25. Gong, Y. and Liu, X., "Creating generic text summaries", *Sixth International Conference on Document Analysis and Recognition*, 9/10-9/13, 2001, pp: 903-907, Seattle, WA, USA.
26. Holsapple, C. W. and Joshi, K.D., "Description and analysis of existing knowledge management frameworks", in *Proceedings of the 32nd Hawaii International Conference on System Sciences* (15) 1999, Los Alamitos, CA, USA.
27. Knight, K. and Marcu, D., "Summarization beyond sentence extraction: a probabilistic approach to sentence compression", *Artificial Intelligence* (139) 2002, pp: 91-107.
28. Lam. W. and Ho, K. S., "FIDS: an intelligent financial Web news articles digest systems", *IEEE Transactions on Systems, Man and Cybernetics-Part A: Systems and Humans* (31: 6) 2001, pp: 753-762.
29. Lehmam, A., "Text structuration leading to an automatic summary system: RAFI", *Information Processing and Management* (35), 1999, pp: 181-191.
30. Maybury, M. T., "Generating summaries from event data", *Information Processing & Managemen* (31: 5) 1995, pp: 735-751.

31. Morris, A.G., Kasper, G.M. and Adams, D.A., “The effects and limitations of automated text condensing on reading comprehension performance”, *Information Systems Research*, March 1992, pp: 17-35.
32. Neff, M. S. and Copper, J. W., “ASHRAM: active summarization and markup”, in *Proceedings of the 32nd Hawaii International Conference on System Sciences* 1999, Maui, HI, USA.
33. Nie, J., Briscbois, M. and Ren, X., “On Chinese text retrieval”, *Conference Proceeding of ACM-SIGIR*, August 1996, pp: 225-233, Zurich.
34. Rothkegel, A., “Abstracting from the perspective of text production”, *Information Processing & Management* (31: 5) 1995, pp: 777-784.
35. Salton, G., *Automatic Text Processing*, *Addison-Wesley Publishing Company*, New York, 1989, pp: 328-338.
36. Salton, G., Singhal, A., Mitra, M. and Buckley, C. “Automatic text structuring and summarization.” *Information Processing & Management* (33:2) 1997, pp: 193-207.
37. Singhal, A., Salton, G., Mitra, M. and Buckley, C. "Document length normalization," *Information Processing & Management* (32) 1996, pp: 619-633.
38. Sproat, R. and Shih, C., “A statistical method for finding word boundaries in Chinese text”, in *Processing of Chinese and Oriental Languages* 1990, pp: 336-351.
39. Teece, D. J., “Capturing value from knowledge assets: the new economy, markets for know-how, and intangible assets,” *California Management Review* (40:3) 1998, pp: 55-79.
40. Zhi, Z., Hin, H. K. P., Gay, R. K. L, Lin, G. W., and Yang, L. S., “iTSum: one agent-based system for automated text summarizing”, *International Conference on Information-Technology and Information–Network* (3) 2001, pp: 18-25, Beijing, China.