

行政院國家科學委員會補助專題研究計畫成果報告

檢索問題、檢索詞彙及心智模型對資訊尋求和 資訊檢索之影響研究 III

A Study of Information Seeking and Retrieving— from Questions, Search Terms, and Mental Models' Perspectives III

計畫類別：個別型計畫 整合型計畫

計畫編號：NSC89 - 2413 - H - 002 - 019 -

執行期間： 88 年 8 月 1 日至 89 年 7 月 31 日

計畫主持人：黃慕萱

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

執行單位：國立臺灣大學圖書資訊學系

中 華 民 國 89 年 7 月 31 日

行政院國家科學委員會專題研究計畫成果報告

檢索問題、檢索詞彙及心智模型對資訊尋求和資訊檢索之影響研究 III

A Study of Information Seeking and Retrieving—from Questions, Search Terms, and Mental Models' Perspectives III

計畫編號：NSC 89-2413-H-002-019

執行期限：88 年 8 月 1 日至 89 年 7 月 31 日

主持人：黃慕萱 國立臺灣大學圖書資訊學系

E-Mail：mhhuang@ccms.ntu.edu.tw

研究助理：陳明君 E-Mail：ivy@ms.cc.ntu.edu.tw

一、中文摘要

本研究分析光碟資料庫資訊需求者和計畫檢索者問題認知、檢索概念、檢索詞彙、詞彙選用心智模型及檢索結果之差異，以探討檢索問題、檢索詞彙以及心智模型對資訊尋求和資訊檢索之影響。本計畫蒐集 10 位資訊需求者之檢索問題，經由 31 位計畫檢索者，對此 10 個指定的問題進行 59 次檢索。本研究將從問題的複雜性、專指性和前提假定三方面分析資訊需求者與計畫檢索者問題認知上之差異，同時分析兩者對於檢索概念、檢索詞彙、詞彙選用心智模型與檢索結果的一致性，以建立使用者導向資訊尋求行為的本土實證研究資料，期能將研究結果應用於資訊系統的設計與訓練上。

關鍵詞：資訊需求、資訊尋求、資訊檢索、問題、使用者、檢索詞彙、檢索概念、心智模型、一致性

Abstract

This project analyzes the information seeking and information retrieval of college and graduate students in National Taiwan University when they are using PsycLIT database. This project collects 10 real search questions from end-users. After instructions, 31 project searchers were assigned to search the above 10 questions, doing 59 searches. This study compares the differences between end-users and project searchers concerning search concepts, search terms, and mental models. It also attempts to establish relationships between questions and the above factors.

Keywords: Information Need, Information Seeking, Information Retrieval, Questions, Users, Search Terms, Search Concepts, Mental Models, Consistency

二、緣由與目的

本計畫是一為期三年的大型實證研究，重點在於研究檢索問題、檢索詞彙及心智模型對資訊尋求和資訊檢索之影響。本計畫已於第一年度和第二年度計畫中分別分析資訊需求者和計畫檢索者之檢索問題認知、檢索概念、檢索詞彙、詞彙選用心智模型以及檢索結果，至於本年度的研究重點在透過比較資訊需求者與計畫檢索者對於檢索問題認知、檢索概念、檢索詞彙、詞彙選用心智模型及檢索結果之差異，探討資訊需求在檢索行為中的角色，並綜合歸納檢索問題、檢索詞彙與個人心智模型對資訊尋求和資訊檢索之影響。

本計畫欲達成之研究目的如下：

- 一、探討資訊需求者與計畫檢索者對於檢索問題認知之差異。
- 二、探討資訊需求者與計畫檢索者檢索概念之差異。
- 三、探討資訊需求者與計畫檢索者檢索詞彙之差異。
- 四、探討資訊需求者與計畫檢索者詞彙選用心智模型之差異。
- 五、探討資訊需求者與計畫檢索者檢索結果之差異。
- 六、綜合歸納檢索問題、檢索詞彙與心智模型對資訊尋求和資訊檢索之影響。

三、結果與討論

本研究彙整所指定之 10 個檢索問題的資訊需求者和計畫檢索者對問題認知、檢索概念、檢索詞彙、詞彙選用心智模型及檢索結果等資料，進行系統化的分析比較，所得結果討論如下：

1. 資訊需求者和計畫檢索者對於檢索問題認知之差異

資訊需求者和計畫檢索者對於檢索問題認知差異的分析，可分為複雜性、專指性和前提假定三

方面：在複雜性方面，有八成的計畫檢索者對問題複雜性認知的平均值高於資訊需求者，原因在於計畫檢索者對於問題的掌握程度不及資訊需求者，對問題和資訊需求者的需求無法深入瞭解，比較容易將問題複雜化，此與相關研究中次判斷者所進行的相關判斷較為鬆散有異曲同工之妙。深入探討各問題之資訊需求者與計畫檢索者對問題複雜性認知之一致性，發現一致性達八成以上的只有一個檢索問題，另有 6 個檢索問題之各計畫檢索者對於問題複雜性認知與資訊需求者完全不一致（共佔 60%），而其它 3 個問題的一致性亦不高。至於一致性較高的 2 個問題皆具備陳述較為簡單而明確，且不包含前提假定的特質，可知需求者對於問題陳述的愈明確，甚至說出其檢索概念和詞彙時，則模擬的計畫檢索者愈能提高對於問題本身掌控程度。

在專指性方面，將各計畫檢索者對問題專指度認知與資訊需求者相比，發現有 39%（21 位）的計畫檢索者之專指度較資訊需求者低。再整合從各問題來看，10 個問題中有 4 個問題之計畫檢索者問題專指性平均值低於該資訊需求者之分析。而差距最小的兩個問題在語意和語法上皆十分清楚，故其專指性平均值差距較小。

在前提假定方面，本研究所選擇的 10 個檢索問題中，有 5 個問題包含前提假定，研究結果發現當問題中以「X 在 Y 上之應用」以及「X 和 Y 之關係」來表達其資訊需求時，不論資訊需求者或是各計畫檢索者皆不將此前提假定直接用來檢索，而將數個概念用交集的方式進行檢索。至於問題中包含與「時間」相關的前提假定時，計畫檢索者對問題陳述中之“近期”和“1980 年以後”認知不同，造成採取的檢索策略不同，因此檢索結果便有相當差異。同時也發現當問題陳述中有「X 對 Y 之影響」方面的前提假定時，計畫檢索者會直接將“影響”用來檢索，但是資訊需求者是將數個概念用交集的方式進行檢索，檢索結果的差異便由此產生。此外，研究發現與資訊需求者比較，計畫檢索者由於處於較不確定的情況下，對於問題的掌控程度較低，因此傾向將前提假定直接用來檢索。

綜合來看，不論從問題之複雜性、專指性或前提假定進行比較，可發現個人對於問題認知之有很大的差異，而此認知上之差異便對於資訊尋求行為造成影響。此外，可歸納出由於計畫檢索者處於較不確定的情況下，對於問題掌握程度不及需求者，因此對於問題認知較為複雜，同時傾向直接將前提假定用來檢索。至於專指性方面，計畫檢索者則略高於資訊需求者。

2. 資訊需求者和計畫檢索者檢索概念之差異

研究資訊需求者與計畫檢索者針對相同問題檢索時，檢索概念的一致性，發現一致性較高的問題，具備陳述十分清楚，且將關鍵的查詢詞彙明確地表達出來之特質，因此計畫檢索者在模擬檢索時，很容易掌握住檢索的中心概念。此外，分三個層次分析資訊需求者和計畫檢索者檢索概念之一致性，發現在資訊需求者與計畫檢索者檢索概念一致性上，從第一層次到第三層次，由於比對標準的

漸次放寬，計算出的一致性呈現提高的現象，但是並沒有任一計畫檢索者與該問題之資訊需求者所列之檢索概念完全相同，與 Saracevic 和 Baxter 的研究結果相反（其研究結果發現不同檢索者之檢索概念一致性高達 78%），而儘管比對層次放寬到第三層次，一致性達六成以上的仍不多，僅佔 26.7%（14 組），可見兩者之檢索概念差異很大。

3. 資訊需求者和計畫檢索者檢索詞彙之差異

在資訊需求者和計畫檢索者檢索詞彙研究方面，首先說明各個問題之資訊需求者與計畫檢索者採用的檢索詞彙數之分析結果，發現 10 個問題中有 8 個問題之計畫檢索者平均使用的檢索詞彙數皆多於該問題之資訊需求者，原因在於資訊需求者對題目較能掌握，而計畫檢索者對於資訊需求的瞭解程度有限，因此傾向使用較多檢索詞彙嘗試檢索。再說明資訊需求者與計畫檢索者檢索詞彙一致性的分析結果，發現檢索詞彙選用一致性比對由第一到第三層次，並無任一計畫檢索者與該問題之資訊需求者的檢索詞彙完全一樣，但由於比對標準的漸次放寬，計出的一致性呈現增高的現象。然而儘管已將比對標準放寬至兩詞彙間有廣狹義與相關關係者皆視為一致，亦只有一成（6 組，佔 10.2%）的計畫檢索者與其資訊需求者採用一致的檢索詞彙，可見兩者的差異極大，此與 Saracevic 和 Baxter 的研究結果（不同檢索者間之檢索詞彙差異很大）相同。

4. 資訊需求者和計畫檢索者詞彙選用心智模型之差異

在資訊需求者和計畫檢索者詞彙選用心智模型方面，在 10 個問題中，一致性較高的問題有 2 個，一致性皆為 33.33%；一致性為 20% 的有 1 個，一致性為 16.67% 者有 4 個，至於其它 2 個檢索問題之各計畫檢索者對於詞彙選用的心智模型皆與其資訊需求者完全不一致。詳細分析一致性較高的 2 個檢索問題，發現其資訊需求者和計畫檢索者之檢索概念及檢索詞彙一致性皆低，但其詞彙選用的心智模型一致性偏高。事實上，歸納本研究結果，發現各計畫檢索者與其問題需求者詞彙選用的心智模型一致性整體來說並不高（最高者為 33.33%），也就是說個人差異在詞彙選用的心智模型上有相當的影響，而此差異便對於資訊尋求行為造成影響。

5. 資訊需求者和計畫檢索者檢索結果之差異

檢索結果分為檢索行數、檢出筆數、相關筆數和檢索精確率來看，先就檢索行數而言，有 74%（37 位）之計畫檢索者的檢索行數少於資訊需求者，若從檢索問題來看，10 個問題中有 8 個問題之各計畫檢索者檢索行數的平均值皆少於該問題之資訊需求者，由此可知，資訊需求者在進行檢索時，較努力嘗試從各個角度進行檢索，以滿足其資訊需求。再比較資訊需求者和計畫檢索者的檢出筆數，發現 54%（共 27 位）的計畫檢索者所檢出的筆數少於資訊需求者，應是由於計畫檢索者對於問

題的不確定較大，因此檢索到的書目較少。而相關筆數方面，有 62%（共 31 位）的計畫檢索者所判斷的相關書目少於該問題之需求者，若以問題來看，有 5 個問題之各計畫檢索者判斷相關的筆數平均值明顯少於需求者，此 5 個問題的複雜性較高、專指性較低，可能由於如此，計畫檢索者進行相關判斷時較困難，因此判斷相關的書目筆數少於資訊需求者。此外，再就精確率來看，有 62%（共 31 位）的計畫檢索者之檢索精確率低於資訊需求者。同時發現當問題較為複雜且專指時，計畫檢索者的檢索精確率才會較資訊需求者高。

深入分析資訊需求者和計畫檢索者之檢索結果的重覆性，發現有 84%（共 42 次）之檢索結果重覆性在 0 和 4.99% 間，其中 37 次（74%）檢索之檢出書目是完全不重覆的，可見檢出書目的重覆性極低。而有 5 次檢索之檢索結果經比對之後重覆性在 5% 和 9.99% 間（佔 10%），有 2 次檢索之重覆性為 35% 至 39.99% 間（佔 4%），重覆性最高的是 43%（僅佔 2%）。從檢索結果分析來看，儘管計畫檢索者與資訊需求者具備相同的學科背景，同時模擬需求者的檢索行為，但是檢索到的書目重覆性極低，有七成以上是完全不重覆的，與 Saracevic 和 Kantor 的研究結果：不同檢索者之檢索結果重覆性不高相同。

6. 檢索問題 檢索詞彙與心智模型對資訊尋求和資訊檢索之影響

在本研究結果中，發現資訊需求者和計畫檢索者不論在檢索問題認知、檢索概念、檢索結果、檢索詞彙和詞彙選用心智模型上的差異皆很大，但是仍能從異中求同，而這些相同點則是系統設計時的重要參考：

綜合歸納檢索問題、檢索詞彙與心智模型對資訊尋求和資訊檢索之影響，可知當資訊需求者和計畫檢索者對問題認知一致性高時，則選用之檢索詞彙一致性高；而當兩者的檢索概念一致性高時，則檢索結果重覆性高。至於當檢索概念一致性高時，則選用檢索詞彙一致性高。此外，由於計畫檢索者處於較不確定的情況下，較無法充分地掌握檢索問題，因此對於問題認知較為複雜，但是專指性較高，且傾向直接將前提假定用來檢索。同時分析資訊需求者與計畫檢索者檢索概念及檢索詞彙之差異，發現兩者之檢索概念和檢索詞彙並沒有完全一致的，但是彼此間有單複數、同義、廣狹義及相關等關係。而當需求者之問題陳述十分清楚，尤其是將關鍵的查詢詞彙明確地表達出來時，計畫檢索者較容易掌握住檢索的中心概念，因此其檢索概念及檢索詞彙與資訊需求者較一致。

深入探討檢索問題認知、檢索概念、檢索詞彙、詞彙選用心智模型和檢索結果一致性間的關係，從一連串有系統的分析，可發現其中有 3 個檢索問題的複雜性高且專指性低，而此 3 個問題之資訊需求者和計畫檢索者之檢索概念、檢索詞彙及檢索結果的一致性皆較低。詞彙選用的心智模型一致性較高，也就是說當問題的複雜性高、專指性低時，資訊需求者和計畫檢索者之檢索概念、檢索詞彙和檢索結果一致性低，而詞彙選用心智模型一致

性高。

四、計畫成果自評

目前在國內非常缺乏有關資訊尋求與資訊檢索的大型實證研究。由於人類行為受語言、社會文化及教育等影響甚鉅，因此很難將歐美各國的研究成果，全盤移植到國內的系統設計上，所以更凸顯以國人為實際研究對象之研究的重要性。而本計畫是以國人為研究對象，並從使用者的角度系統化地研究國人的資訊檢索過程及資訊尋求行為，其研究結果將是設計高品質資訊系統的最佳參考依據。此外，因西文資料庫檢索在可預見的將來仍將持續作為國人資訊檢索的主流，本計畫亦提供一研究國人於非母語資訊系統之資訊尋求行為的絕佳機會，其結果可作為西文資料庫引進國內時，界面修改或利用指導的參考。

五、參考文獻

- [1] Iivonen, Mirja. "Consistency in the Selection of Search Concepts and Search Terms." *Information Processing and Management* 31:2 (March-April 1995): 173-190.
- [2] Iivonen, Mirja. "From Translation to Navigation of Different Discourses: A Model of Search Term Selection During the Pre-Online Stage of the Search Process." *Journal of the American Society for Information Science* 49:4 (April 1998): 312-326.
- [3] Saracevic, Tefko and Matthew A. Baxter. "Measuring the Degree of Agreement Between Searchers." *Proceedings of the 47th Annual Meeting of the American Society for Information Science* v.21 (1984): 227-230.
- [4] Saracevic, Tefko and Matthew A. Baxter. "On a Method for Studying the Structure and Nature of Requests in Information Retrieval." *Proceedings of the 46th Annual Meeting of the American Society for Information Science* 20 (1983): 22-25.
- [5] Spink, Amanda and Tefko Saracevic. "Interaction in Information Retrieval: Selection and Effectiveness of Search Terms." *Journal of the American Society for Information Science* 48:8 (August 1997): 741-761.
- [6] Spink, Amanda. "Term Relevance Feedback and Mediated Database Searching: Implications for Information Retrieval Practice and Systems Design." *Information Processing & Management* 31:2 (March-April 1995): 161-171.