

圖書分類自動化 原理誘導에 관한 研究

沈 宜 順* · 李 慶 浩**

〈目 次〉

- | | |
|---------------------|---------------------------|
| I. 緒 論 | 3. CC의 基本原理 |
| 1. 研究目的 | IV. 分類自動化시스템을 위한 原理
誘導 |
| 2. 研究方法 | 1. 시스템의 概要 |
| 3. 研究의 範圍 및 限界 | 2. 用語蒐集에 따른 코드화方法 |
| II. 分類自動化에 관한 史的 概觀 | 3. 分類用 파일設計 |
| 1. 分類表的 側面에서의 研究 | 4. 分類自動化的 基本原理 |
| 2. 情報檢索의 側面에서의 研究 | V. 適用上の 問題點 및 提言 |
| III. 롤론分類法의 基本原理 | VI. 結 論 |
| 1. CC의 構成 | |
| 2. CC의 分類體系 | |

I. 緒 論

1. 研究目的

圖書館 및 情報센터 등에 있어서 分類는 文獻의 內容을 概念的으로 分析하고 文獻에 대한 効率的인 接近을 할 수 있도록 組織하는 過程이다. 그러나 이것은 複雜한 內容을 지니고 있는 主題를 그 構成要素인 重要概念으로 分析하고 組織하는 主題分析의 過程을 要하기 때문에 相當한 人的, 經濟的 부담을 안겨주고 있다.

특히 20C 以後에 접어들면서 부터 기하급수적으로 增加하고 있는 文獻生産으로 인하여 大規模의 圖書館들은 漸進的으로 自動化를 通하여 利用者에

* 大邱大學校 圖書館學科 助教授

** 大邱大學校 圖書館學科 專任講師

2 도서관학논집

제 보다 迅速하고 正確한 奉仕를 遂行하고 있으나, 分類業務만은 아직까지도 文獻을 書架上에 配架하기 위한 分類過程의 自動化, 즉 코드화 分類의 自動化는 他 業務에 비하여 크게 낙후된 狀況에 처하여 있다고 하겠다.

물론 圖書館이 自動化된 경우에는 그 文獻의 內容을 代表할 수 있는 키워드(keywords)를 비롯하여 著者名, 書名, 分類記號, ISBN(International Standard Book Number) 등의 多様な 接近方法이 實現되어 질 수 있기 때문에 分類를 하지 않고 收入順으로 文獻을 書架上에 配架하는 方法도 있을 수 있다. 그러나 이러한 경우에는 文獻檢索은 컴퓨터를 通하여 迅速히 이루어 질 수 있으나, 檢索된 個個文獻들에 대한 書架上의 接近은 같은 主題의 文獻이라 하더라도 書架別, 書庫의 層別로 分離配架되기 때문에 文獻接近時, 특히 主題別 關聯文獻接近의 경우에는 매우 非効率的이라 하겠다.

이에 本 研究는 未來의 圖書館들이 漸進적으로 自動化될 경우를 考慮하여 오늘날 大部分의 圖書館에서 使用하고 있는 列擧式 分類方式과는 달리 概念分析的 接近方法에 의한 分析合成式的 原理로 이루어진 콜론分類法(Colon classification)의 原理를 應用하여, 分類의 自動化가 可能할 뿐만 아니라, 圖書館과 利用者의 側面에서 檢索과 利用이 便利한 分類自動化를 위한 하나의 原理를 誘導하여 보고자 하는 데 있다.

2. 研究方法

本 研究는 分類의 自動化를 위한 原理誘導에 관한 하나의 方法論을 提示하고자 한다. 따라서 오늘날 많은 圖書館에서 使用하고 있는 DDC, KDC, LC 등과 같은 列擧式 分類方法과는 달리 概念分析的 接近方法에 입각한 CC의 原理를 應用하여 自動적으로 分類記號가 生成되어 질 수 있도록 하는 分類自動化를 위한 一般的인 原理誘導를 위한 것으로 이의 研究方法를 보면 다음과 같다.

(1) 콜론分類法의 分類體系를 分析하여 CC의 基本原理를 밝혀낸 後, 이로부터 分類自動化가 可能한 一般的 原理 및 應用原理를 誘導하였다.

(2) 이 分類自動化原理에 의한 自動分類를 위한 分類用과일 設計方法을 說明하였다.

(3) 分類用과일은 特定主題分野만을 위한 경우와 一般分類用과일 編成의 두가지 側面으로 나누어 設計하였다.

(4) 實際 分類上에서 發生하는 諸 問題點을 檢討하여 보았다.

(5) 本 研究上에서 提示한 分類自動化原理에 의한 適用可能分野와 適用不可能分野 및 適用上에서 發生하는 諸 問題點을 檢討하여 보았다.

3. 研究의 範圍 및 限界

本 研究의 範圍 및 限界는 다음과 같다.

(1) 本 研究는 코드分類의 自動化를 위한 方法論 및 一般的인 原理만을 다루고자 한다.

(2) 分類는 文獻의 標題 혹은 키워드를 컴퓨터에 入力하여 完全自動化에 의거 分類하고자 한다.

(3) 分類의 基本原理는 CC에 의거한다.

(4) 分類原理의 適用對象은 全學問分野에 適用可能한 一般的 原理를 誘導하고자 하나 모든 學問分野에 完全히 適用可能하지는 않을 것으로 思料되는 바 이들 學問分野에 대하여는 그 學問分野의 特殊性에 맞는 應用原理를 適用하고자 한다.

(5) 實際 分類過程의 實驗은 행하지 않고 理論的 側面에서만 다룬다.

(6) 本 研究는 分類自動化에 대한 계속적인 研究의 一部分임을 밝혀 두고자 한다.

II. 分類自動化에 관한 史的 概觀

圖書館業務에 대한 컴퓨터處理의 最初 試圖는 1954年 美國海軍兵器試驗所

(海軍兵器센타)의 附屬圖書館⁽¹⁾이다.

그러나 分類의 自動化에 있어서는 이보다 훨씬 뒤인 1960年代에 접어들면서부터 情報檢索을 위한 文獻分類(document classification)의 自動化와 分類表自體의 研究를 通한 圖書分類(코드分類)의 自動化와 같은 兩大側面으로 그 研究가 活潑히 進行되어 온 바, 이에 대한 史的 概觀을 살펴보면 다음과 같다.

1. 分類表的 側面에서의 研究

分類表的 側面에서의 研究는 주로 分類記號를 부여하는 코드分類的 形態로서 이에 대한 研究를 ① 1950年代 以前, ② 1950年代, ③ 1960年代, ④ 1970年代 등으로 區分하여 보면 다음과 같다.

가. 1950年代 以前

1937年 프랑스 파리에서 開催된 世界도큐멘테이션會議(World Conference of Documentation)에서 그로리어(Grolier)가 包括的 分類(comprehensive classification)을 위한 一般的 原理 및 理論을 診斷하여 보도록 強調⁽²⁾한 事實은 하나의 學問으로서 分類에 관한 研究의 始發이라 할 수 있다. 또 1948年 英國學士院의 科學情報會議(Royal Society's Scientific Information Conference)⁽³⁾에서 討議된 主題의 하나가 分類에 관한 것이었으며, 이 會議의 結果로 分類研究를 위한 委員會가 發足하게 되고, 베르날(J. B. Bernal)教授가 이 委員會의 業務를 擔當하기에 이르러 分類에 대한 研究가 進行되기에 이른 時期이다.

나. 1950年代

1950年代에 접어들면서 分類에 대한 關心은 점차 고조된다. FID에 대한

(1) Frederick G. Kilgour, "History of Library Computerization," *Journal of Library Automation*, Vol. 3, No. 3(1970), p. 161.

(2) P. S. G. Kumer, "International Study Conference on Classification Research(3)," *Herald of Library Science*, Vol. 14, No. 1 (January 1975), p. 8.

(3) A. C. Foskett, *The Subject Approach to Information*. 3rd ed. (London: Clive Bingley, 1977), p. 51.

란가나단(S. R. Ranganathan)의 끈질긴 努力과, 당시 FID의 事務局長 (Secretary General) 쿠퍼스(Donker Duyvis)의 구상으로 1950年 FID/CA (=Committee on General Theory of Classification)가 發足되고 란가나 단이 Rapporteur General이 되었다.⁽⁴⁾ 이로 인하여 分類도 단순한 學問 (discipline)에서 뿐만 아니라 하나의 科學(science)으로서 國際的 次元에서 研究하고자 하였다.

한편 베르날 教授가 擔當하여 오던 分類委員會의 業務는 1951年 비커리 (B. C. Vickery)教授가 引繼받아 1952年 CRG(Classification Research Group)⁽⁵⁾라는 公式名稱으로 바꾸어 發足시켰다.

이 CRG는 1952年 웰즈(A. J. Wells), 비커리, 코우즈(E. J. Coates), 페라던(J. Farrodane), 포스켓(D. J. Foskett), 밀즈(J. Mills) 및 팔머 (B. E. Palmer)⁽⁶⁾등을 構成員으로 하여 그 最初의 會議을 開催하였다. 이들은 이 會議을 通하여 既存의 分類表에 대한 不滿을 나타냄과 동시에 독특한 分類方法을 模索하려고 하였다. 이에 1953年 파세트(facet)分類에 대한 自身들의 見解를 밝힘과 동시에 1955년에는 “情報檢索의 모든 方法의 基礎는 파세트分類”란 비망록을 刊行하였다. 이 비망록에는 ① 圖書分類의 基礎로서 파세트分類, ② 페라던의 카테고리間的 關係에 관한 理論, ③ 單純한 記號法(notation)⁽⁷⁾의 使用 등 세가지 基本觀念(basic ideas)을 強調하고 있다.

1957年 第1回 情報檢索을 위한 國際分類研究會議(International Study Conference on Classification for Information Retrieval=ISCCR-I)가 CRG와 런던大學 圖書館學校(School of Librarianship and Archives)의 協助로 Aslib(Association of Special Libraries and Information Bureau)에 의하여 도킹(Dorking)에서 開催⁽⁸⁾되었는 바, 이 會議의 結果로 圖書館

(4) P. S. G. Kumer, op. cit., p. 9.

(5) A. C. Foskett, op. cit., p. 51.

(6) Krishan Kumar, *Theory of Classification* (New Delhi: Vikas Publishing House, c1979), p. 466.

(7) Ibid., p. 466.

(8) E. J. Coates, "Classification in Information Retrieval; the Twenty Years following Dorking," *Journal of Documentation*, Vol. 34, No. 4 (December 1978), p. 289.

專門人 및 情報專門家가 한자리에 모여 分類에 관한 새로운 見解 및 方法論을 討議할 轉換點을 마련함과 동시에 파세트分類가 모든 情報檢索의 基礎가 된다고 宣言⁽⁹⁾하였다.

다. 1960年代

1950년에 發足한 FID/CA는 1961年 FID/CR(=Committee on Classification Research)로 再編成한 뒤, 1964年 덴마크의 엘시노어(Elsinore)에서 第2回 國際分類研究會議(ISCCR-2)을 開催하였는데,⁽¹⁰⁾ 이 會議의 主要 議題가 分類의 一般理論과 自動化分類였다. 특히 FID/CR은 62個國의 FID 멤버로 構成되었기 때문에 國際의 次元에서 가장 영향력 있는 集團이었으며, 個個멤버들도 FID/CR의 業務에 積極的인 자세로 임하였다.⁽¹¹⁾

한편 1960年代에 접어들면서 CRG는 特殊分類表와 一般分類表(general classification) 사이의 關係 및 파세트原理에 의한 새로운 一般分類表(new general classification)의 編纂과 같은 分野에 關心의 焦點을 두었으며, 1962년에는 새로운 一般分類表의 開發을 위한 NATO 寄金⁽¹²⁾이 英國圖書館協會에 提供되자, 協會側은 이 寄金을 다시 CRG에게 委任하였다. 이에 CRG의 톰린슨(H. Tomlinson)이 1964~1968년까지, 오스틴(Derek Austin)이 1968~1969년까지 이 計劃에 參與하였는데, 이 때 CRG는 ① 概念의 카테고리화를 위한 原理決定(determining principles for the categorization of concepts), ② 카테고리內에서의 概念의 順序(ordering concepts within categories), ③ 概念사이의 關係(relationships between concepts) 등에 研究를 置重한 바, 이 期間中에 研究되어진 CRG의 圖書館學關係分類表(Classification of Library Science)⁽¹³⁾는 파세트原理에 의거한 代表的인 分類表라고 하겠다.

(9) A. C. Foskett, op. cit., p. 204.

(10) P. S. G. Kumer, op. cit., p. 9.

(11) Krishan Kumar, op. cit., p. 467.

(12) Ibid., p. 468.

(13) Classification Research Group, "Bulletin no. 9," *Journal of Documentation* vol. 24, No. 4 (December 1968), pp. 290-292.

라. 1970年代

1970年代에 접어들면서 CRG는 書架上的 配列 및 情報檢索의 機能을 함께 지닐 수 있는 兩立分類表(compatible scheme)의 考案을 目標로 하였었다. 그러나 이 두가지 機能의 兩立은 不可能하다는 事實이 立證되자⁽¹⁴⁾ 自動檢索시스템을 위한 分類表의 開發로 그 方向을 轉換하였다. 이러한 狀況下에서 과제분류에 입각한 自動化시스템을 위한 最初의 所産物이 바로 PRECIS(PREserved Context-Indexing System)이다.

그리고 第3回 國際分類研究會議(ISCCR-3)가 1975年 1月 6日에서 11日까지 印度의 봄베이에서 開催되었는데, 이 會議의 主題가 全世界의 情報網을 위한 分類시스템(Ordering Systems for Global Information Networks)⁽¹⁵⁾ 이었다. 그러나 이것은 당초 UNISIST(World Science Information System)가 情報의 流進을 圓滑하게 하기 위하여 標準參照코드(Standard Reference Code)라 불리는 一種의 코드화에 의한 分類인 標準網羅分類表(Standard Roof Classification)가 開發되어야 한다는 決定에 따라 FID의 로이드(G. A. Lloyd) 및 코우즈(C. A. Coates)등이 이를 위탁받아 새로운 分類表의 開發을 試圖한 것인데 이 SRC가 後에 廣域分類시스템(Broad System of Odering)으로 合併되어진 것이다.

2. 情報檢索의 側面에서의 研究

情報檢索의 側面에서의 分類自動化에 관한 研究는 주로 用語分類(keyword classification)와 文獻分類(document classification)⁽¹⁶⁾이다. 用語分類는 情報檢索時 利用者の 質問과 文獻의 內容을 나타내는 用語가 서로 일치하는 範圍를 넓히기 위하여 個個의 用語를 類似語群으로 分類시키는 方法이고 文獻分類는 서로 關聯性있는 文獻을 서로 群集化(grouping)시켜 檢索時 한

(14) Krishan Kumar, op. cit., pp. 478-479.

(15) P. S. G. Kumer, op. cit., pp. 10-26.

(16) G. Salton, *Dynamic Information and Library Processing* (Englewood Cliffs, NJ: Prentice-Hall, c1975), p. 223.

文獻群단을 探索하도록 制限시켜 檢索의 迅速性を 기하기 위한 方法이다.

情報檢索의 側面에서 分類에 관한 最初의 研究는 1952年 *IBM Journal of Research and Development* 에 發表한 바 있는 論(H. P. Luhn)⁽¹⁷⁾ 이라고 할 수 있다. 이것은 물론 圖書의 書架上 配列을 前提로한 分類는 아니더 文獻에 있어 單語의 統計的 分析이 文獻의 內容分析에 몇가지 竝다리를 提供하여 준다는 事實을 보여 준 것이다. 그러나 分類의 自動化를 앞시한 最初의 사람은 19世紀(1856?) 크레스타도로(Crestadore)⁽¹⁸⁾ 로서 그 당시에 벌써 順列索引技法을 提示하였다.

文獻檢索의 側面에서 分類의 自動化에 관한 研究가 활발해 지기는 1966年 代에 접어들면서 부터로 그 代表的인 研究者들을 보면, 劍브리지 言語研究 所의 스파크 존스(K. Sparck Jones)와 니드햄(R. M. Needham),⁽¹⁹⁾ 스파크 존스와 잭슨(D. Jackson),⁽²⁰⁾ 프라이스(N. Price)와 스킴비노치비(S. Schiminovich),⁽²¹⁾ 케슬러(M. M. Kessler),⁽²²⁾ 콕(K. L. Kwok)⁽²³⁾ 살톤(G. Salton)⁽²⁴⁾ 및 보코(H. Borko)와 베르니크(M. Bernick)⁽²⁵⁾ 등의 研究를 들 수 있다.

(17) Donald V. Black, "Automatic Classification and Indexing for Libraries," *Library Resources & Technical Services*, Vol. 9, No. 1 (Winter 1965), p. 36.

(18) John Metcalfe, *Information Retrieval, British & America 1876-1976* (Metuchen: The Scarecrow Press, Inc., 1976), p. 34.

(19) K. Sparck Jones and R. M. Needham, "Automatic Term Classification and Retrieval," *Information Storage and Retrieval*, Vol. 4, No. 2 (1968), pp. 91-100.

(20) Karen Sparck Jones and David Jackson, "Current Approach to Classification and Clump-Finding..." *Computer Journal*, Vol. 10, No. 5 (May 1967), pp. 26-37.

(21) Nancy Price and Samuel Schiminovich, "A Clustering Experiment: First Step Towards a Computer-Generated Classification Scheme," *Information Storage and Retrieval*, Vol. 4 (1968), pp. 271-280.

(22) M. M. Kessler, "Bibliographic Coupling between Scientific Papers," *American Documentation*, Vol. 14, No. 1 (Jan. 1963), pp. 10-11.

(23) K. L. Kwok, "The Use of Title and Cited Titles as Document Presentation for Automatic Classification," *Information Processing & Management*, Vol. 11 (1975), pp. 201-206.

(24) Genard Salton, op. cit.

(25) Harold Borko and Myrna Bernick, "Automatic Document Classification," *Journal of Association for Computing Machinery*, Vol. 10, No. 2 (April 1963), pp. 151-161.

Ⅲ. 콜론分類法の 基本原理

1. CC의 構成

CC를 創案해 낸 란가나단은 數學教師였다. 그러나 1924年 印度 Madras 大學의 圖書館司書로 임명됨과 동시에 英國의 런던大學에서 圖書館學을 공부하기 위하여 英國으로 留學하였다. 이 때 란가나단은 既存의 圖書館分類體系에 불만을 갖고 새로운 分類法을 考案하고자 하였다.

英國에서 修學 後, 1925年 6月 母國 印度로 歸國中에 배안에서 英國런던에서 初案하여 든 임시分類表(tentative schedules)를 손질하여 3,000卷의 Madras 大學圖書館 藏書目錄과 船體內的 圖書館藏書를 分類하였다. 1927年에는 分類表가 거의 完成되어 Madras 大學圖書館藏書를 거의 分類한 後, 書架上에 配架하기도 하였고, 이들 分類記號에 대한 利用者들의 反應을 觀察(1928-31) 修定 後 1933년에 初版이 刊行되었다.⁽²⁶⁾ 그 後 1939년에 2版, 1950년에 3版, 1952년에 4版, 1957년에 5版, 1960년에 6版이 刊行되었으며, 現在 7版이 準備中에 있다.

CC는 이렇게 改定될 때 마다 새로운 記號나 facet를 追加시킴과 아울러 facet의 組合公式도 改定하였다. 第6版은 第1篇이 分類規定, 第2篇이 分類表, 第3篇이 古典과 聖書分類로 되어 있다. 이 CC에서 란가나단은 學問의 專門領域을 42個로 區分함과 동시에 各 學問分野에서 研究되어 질 수 있는 現象, 즉 基本카테고리(fundamental categories)를 P[Personality], M[Matter], E[Energy], S[Space]와 T[Time]로 區分하면서 이들의 論理的 組合順序를 PMEST로 定義하였다.

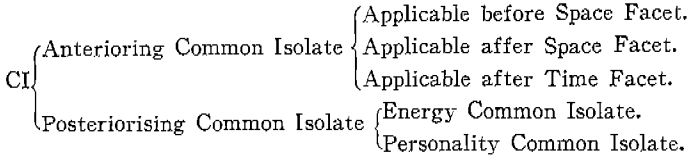
42個 主題分野의 分類에 있어 共通의으로 使用할 수 있는 共通補助分類表로서는 다음과 같은 것이 있다.⁽²⁷⁾

(26) Krishan Kumar, op. cit., pp.52-53.

(27) S. R. Ranganathan, *Colon Classification* (New York: Asia Publishing House, 1966) 제 1, 2권 참고.

가. 共通細目(Common Isolate)

共通細目은 小文字의 이테릭체로 되어 있으며 다음과 같이 區分된다.



나. 時代區分(Time Isolate)

時代區分은 補助表上에 年代區分(Chronological Division)과 Featured Time 으로 區分되어 있다.

다. 地理區分(Space Isolate)

DDC의 地理區分에 해당하는 區分으로서 地球上의 地理的 區分인 Geographical Division 과 地形的인 面을 나타내는 [S2]로 區分되어 있다.

라. 言語區分(Language Isolate)

言語區分은 言學 및 文學類의 分類에 使用되어지는 區分으로서 十進分類法의 言語區分과 같은 性格을 지니고 있다.

마. 各種 Device

上記 列擧한 補助分類表 이외에 괄호속에 들어 表示하는 主題區分(Subject Device), 1~8까지의 숫자를 使用하여 意味를 나타내는 Mnemonic Device, 주로 첫자나 initial을 使用하여 나타내는 Alphabetical Device, 그리고表上에는 나타나 있지 않은 두 isolate를 連結시켜 주는 Superimposition Device, 特殊主題의 強調을 위한 Favored Focus 및 Phrase, Inter-facet, Intra-Array 關係등이 있다.

2. CC의 分類體系

CC의 分類體系는 個個의 主題分野에서 研究되어질 수 있는 主題의 屬性을 PMEST로 區分하고 있으나, 특히 [P]카테고리를 그 主題分野의 特性에 따라 Array시킴과 동시에 그 主題의 屬性에 맞는 概念의 組合이 可能

하도록 하고 있다. 무엇보다도 CC 自體는 概念의 組合에 따른 分類方式을 採擇하고 있기 때문에 그 概念의 屬性에 따라 組合의 方法을 달리하는 것은 明白한 것이라 할 수 있다.

따라서 大部分의 主題分野에서는 바로 PMEST의 組合으로 이루어지나, 그 學問分野의 特性에 따라 Array 시킬 必要性이 있는 分野는 Array 시키고 있는데, 數學, 物理學, 地學, 藥學, Useful Arts, 美術(Fine Arts), 哲學등이 이 部類에 屬한다.

이들 中 物理學(C) 分野를 例로 들어 說明하여 보면, 物理學에서 研究되 어질 수 있는 屬性에 따라 C1 Fundamentals, C2 Properties of Matter, C3 Sound, C4 Heat, C5 Light, Radiation, C6 Electricity, C7 Magnetism, C8 Cosmic hypotheses 등으로 區分시키면서 이들 個個區分들을 物理學內에서 同一한 分類原則을 適用하지 않고 各各 달리 適用하여 分類하고 있다. 이에 物理學分野의 分類表를 圖式化하여 보면 다음 表 1과 같이 나타 낼 수 있다.

또 農學(J)分野의 경우에는 上記 物理學分野가 個個 同立分類體內의 分類項目이 서로 배타적이어서 그 分類 formula도 各各 다른데 비하여 農學分

表 1. 物理學分野 分類表의 構成

Canonical Division	分類 formula	[P1]	[P2]	[E]
Fundamentals: C1	—	C13 Matter C11 Weight C18 Gravitation C14 Energy ⋮	—	—
物 Properties of Matter :C2	C2[P], [P2]	1 Solid 15 Glass 16 Crystal 5 Liquid 8 Gas	13 Density 2 Elasticity 4 Surface } 5 tension } 7 Pressure 76 Vacuum 8 Structure	—

	Canonical Division	分類 formula	[P1]	[P2]	[E]
理	Sound: C3	C3[P]: [E] [2P]	1 Audible Sound 2 Infra Sound 5 Ultra Sound	—	1 Generation 2 Propagation 3 Frequency 5 Interference 7 Acoustics 8 Nature
	Heat: C4	C4: [E] [2P]	—	—	1 Generation 2 Transference 3 Thermometry 4 Exponion of body 7 Thermodynamics
學	Light, Radiation :C5	C5[P]: [E] [2P]	1 Light 2 Ultra violet 3 X 4 Gamma 6 Infra red 7 Hertzian	—	1 Generation 2 Propagation 3 Dispersion 5 Interference 6 Double refraction
	Electricity :C6	C6[P]: [E] [2P]	2 Current 3 Stitical electricity	—	1 Generation 2 Propagation 3 Potential 4 Transformation
	Magnetism :C7	C7[P]: [E]	2 Dia 3 Para 4 Ferro 5 Terrestrial	—	2 Intensity 8 Nature of magnetism
이 하 생 략					

野는 個個 分類項目이 같은 formula 를 갖고 있다. 農學分野의 分類體系를 보면 다음 圖 1⁽²⁸⁾과 같다.

(28) Susan Artandi ed., *Colon Classification* (New Jersey: The Rutgers Univ. Press, c1965), p. 30.

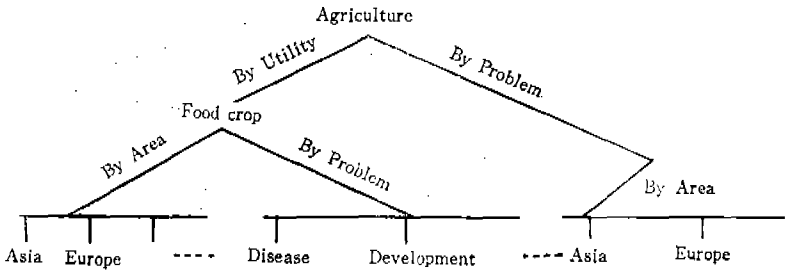
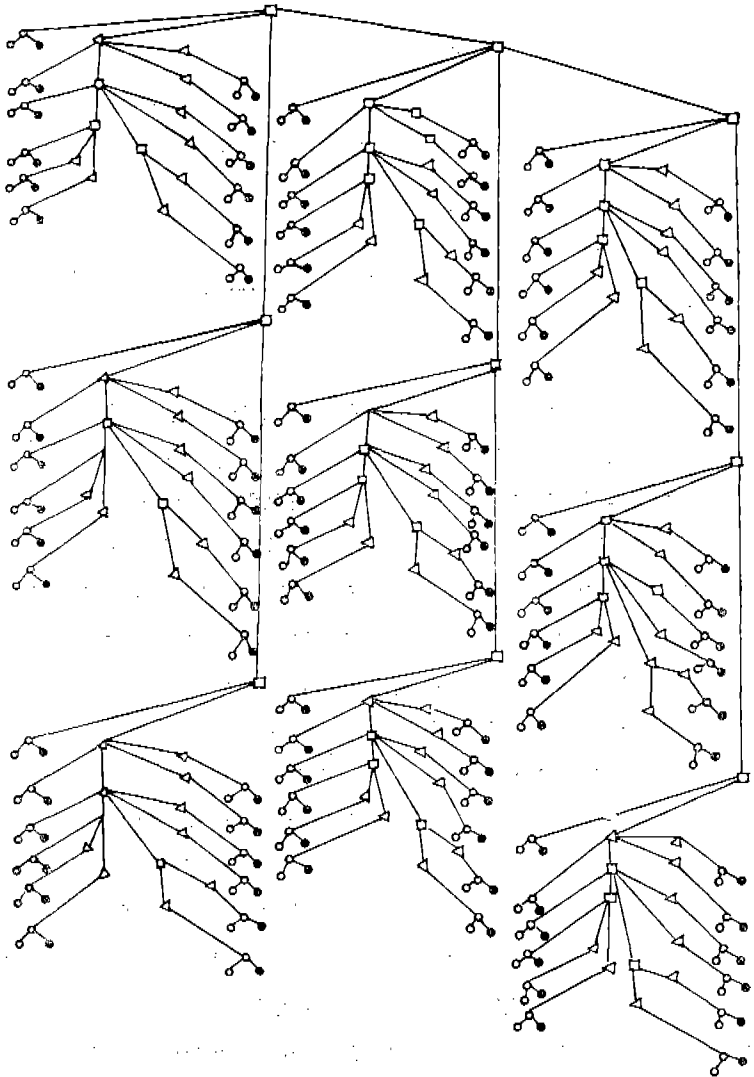


圖 1. 農學分野의 分類體系

圖 1에서 나타낸 바와 같이 農學分野의 경우에는 반드시 [P]카테고리로 分類되는 것은 아니다. 따라서 分類記號의 組合도 農學分野의 特性에 맞게 단일 formula, 즉 $J[P]:[E][2P]:[2E]$ 로 構成되어 있기 때문에 上記公式의 組合인 $J[P]$, $J[P]:[E]$, $J:[E]:[2P]$, $J[P]:[E][2P]:[2E]$ 등으로 組合을 이루면서 分類되어 진다. 또 時代區分이나 地理區分을 適用하여 아시아農業, 유럽農業 및 2000年代의 農業, 1900年代의 農業 등과 같이 [P]카테고리의 組合이 없이도 分類가 자유로이 이루어 질 수 있으며, [P]카테고리와 組合이 되면 그 組合의 數는 더욱 늘어나게 된다.

以上과 같은 原理를 갖고 分類되어 지는 CC의 分類體系를 圖式化하여 보면 다음 圖 2⁽²⁹⁾와 같이 나타낼 수 있으며, 이것은 마치 鍊연관계의 系譜와 같은 모양을 하고 있다. CC는 文獻의 主題를 各 構成要素로 區分하는 方式을 導入하여 이것을 다시 一定의 公式에 따라 組立하는 形態이기 때문에 DDC나 LC와 같이 大部分의 主題에 대하여 이미 分類記號가 정해져 있는 것과는 달리 主題의 構成要素로 되어 있는 基本單位表(standard unit schedules)의 分類 formula에 의하여 分類되는 것이 特徵이다. CC는 圖書分類에 있어 文獻自體가 分類記號를 決定한다는 論理이기 때문에 相對論的이라기 보다는 絶對論이라고 할 수 있으며, 分類記號 自體가 一定의 論理에

(29) Ibid., p.31.



□=Personality Isolate △=Energy Isolate
 ○=Space Isolate ●=Time Isolate
 Matter Isolate is not represented in this diagram.

圖 2. CC의 分類體系

의하여 形成되기 때문에 CC의 體系를 다소 變形하면 컴퓨터에 의한 自動化가 可能한 原理를 誘導하여 낼 수 있다고 하겠다.

3. CC의 基本原理

CC의 基本原理를 간단히 表現한다면, 各 學問分野內에서 그 分野의 特性에 맞게 分析하여 CC表上에 表示되어 있는 [P], [M], [E] 및 [S]와 [T] 카테고리를 組合하여 [P] [M] [E] [S] [T] 順으로 配列하되 [P] 앞에는 ,(콤마), [M] 앞에는 ;(semi-colon), [E] 앞에는 :(colon), [S] 앞에는 .(dot), [T] 앞에는 '(inverted comma)를 부여하는 分類方式이라고 할 수 있다. 그러나 各 學問分野別로 分類 formula가 基本 formula인 PMEST와 다른 경우도 있으나 대체로 다음 表 2에 나타난 바와 같다. 이 表上에 나타난 모든 formula는 앞서 나타낸 바 있는 圖 2의 分類體系를 잘 反映하고 있으며 이것이 바로 CC의 分類體系이자 동시에 CC 分類의 基本原理의 母體가 되는 것이라고 하겠다.

CC의 長點은 바로 分類體系를 圖式化 할 수 있다는 점이며, 分類過程全體가 一定한 루틴(routine)을 갖고 있다는 점이다. 이것이 다른 列舉式 分類方式과 다른 점이며 自動化가 可能한 原理의 定立이 實現될 수 있다는 前提條件이 된다고 하겠다.

表 2. CC의 分類原理

기호	Main Classes	Canonical Divisions		Formula
		기호	Canonical Classes	
a	Generalia Bibliography			a [P], [P2] [P3], [P4]
1	Universe of Knowledge			
2	Library Science			2 [P]; [M]: [E] [2P]
3	Book Science			
4	Journalism			

기호	Main Classes	Canonical Divisions		Formula
		기호	Canonical Classes	
A	Natural Sciences			
β	Mathematical Sciences			
B	Mathematics	B1	Arithmetic	
		B13	Integer	B13 [P], [P2] : [E] [2P]
		B2	Algebra	
		⋮		
		B23	Theory of equation	B23 [P] : [E]
		B25	Higher Algebra	B25 [P], [P2] : [E] [2P]
		B3	Analysis	
		⋮		
		B33	Differential and integral equation	B33 [P], [P2], [P3] : [E] [2P]
		⋮		
		B37	Real variable	B37 [P] : [E] [2P]
		B38	Complex variable	B38 [P] : [E] [2P]
		⋮		
		B6	Geometry	B6 [P] : [E] [2P]
		B7	Mechanics	B7 [P] : [E] [2P]
		B9	Astronomy	B9 [P] : [E] [2P]
C	Physics	C1	Fundamentals	
		C2	Properties of Matter	C2 [P], [P2]
		C3	Sound	C3 [P] : [E] [2P]
		C4	Heat	C4 : [E] [2P]
		C5	Radiation	C5 [P] : [E] [2P]
		C6	Electricity	C6 [P] : [E] [2P]
		C7	Magnetism	C7 [P] : [E]
		C8	Cosmic Hypotheses	
D	Engineering			D [P], [2P] [P2] : [E]
E	Chemistry			E [P] : [E] [2P]
F	Technology			F [P] : [E] [2P]
G	Biology			G [P] : [E] [2P]
H	Geology	H1	Mineralogy	H1 [P] : [E] [2P]
		H2	Petrology	H2 [P] : [E] [2P]

기호	Main Classes	Canonical Divisions		Formula
		기호	Canonical Classes	
		H3 H4 H5 H7	Structural Geology Dynamic Geology Stratigraphy Economic Geology	H7[P]:[E] [2P]
HZ	Mining			HZ[P], [P2]:[E] [2P]
I	Botany			I[P], [P2]:[E] [2P]
J	Agriculture			J[P]:[E]:[2P]:[2E]
K	Zoology			K[P]:[E] [2P]
KZ	Animal Husbandry			KZ[P]:[E] [2P]:[2E] [3P]
L	Medicine			L[P]:[E] [2P]
LE	Pharmacognosy	LZ3 LZ5 LZ8	Pharmacology Pharmacopoeia Pharmacy	LZ3[P]:[E] [2P] LZ5[P], [P2]
M	Useful Arts	M1 ⋮ M7 MJ7	Book production Textiles Rope making	M7[P]:[E] [2P] MJ7[P]:[E] [2P]
△	Spiritual Experience and Mysticism			△[P], [P2]:[E] [2P]
N	Fine Arts	NA NB ND NN NQ NR	Architecture Town planning Sculpture Engraving Painting Music	NA[P], [P2] [P3], [P4]:[E] NB[P], [P2] [P3], [P4]:[E] ND[P], [P2] [P3]; [M]:[E] [2P] NQ[P], [P2] [P3]; [M]:[E] [2P] NR[P], [P2] [P3]; [M]:[E] [2P]
O	Literature			O[P], [P2] [P3], [P4]
P	Linguistics			P[P], [P2] [P3]:[E] [2P]
Q	Religion			Q[P]:[E] [2P]

기호	Main Classes	Canonical Divisions		Formula
		기호	Canonical Classes	
R	Philosophy	R1	logic	R3[P], [P2] R4[P], [P2] R6[P], [P2] [P3], [P4]
		R2	Epistemology	
		R3	Metaphysics	
		R4	Ethics	
		R5	Aesthetics	
		R6	Indian Philosophy	
S	Psychology			S[P]:[E] [2P]
T	Education			T[P]:[E], [2P], [2P2]
U	Geography			U[P]. [S]. [T]
V	History			V[P], [2P]:[E] [2P]. [T]
W	Political Science			W[P], [P2]:[E] [2P]
X	Economics			X[P]:[E] [2P]
Y	Sociology			Y[P]:[E] [2P]:[2E] [3P]
Z	Law			Z[P], [P2], [P3], [P4]

Ⅳ. 分類自動化시스템을 위한 原理誘導

1. 시스템의 概要

分類의 自動化란 컴퓨터의 도움으로 어떤 對象에 대하여 一般性格이나 類似性에 따라 自動적으로 類別로 나누거나 配列하는 行爲를 뜻하며, 圖書分類의 自動化란 컴퓨터의 도움으로 體系적으로 編成된 分類用파일에서 한 圖書의 內容, 主題 또는 形式에 一致하거나 類似한 分類番號를 探索하여 그 圖書에 自動적으로 分類記號를 配定하는 行爲라고 하겠다.

따라서 分類自動化시스템(Automatic Book Classification System)은 우선 分類를 위한 分類用파일을 만들어야 한다. 이 分類用파일이 만들어진 後

에야 비로소 分類가 행하여 질 수 있기 때문이다. 分類는 文獻의 標題를 바로 컴퓨터에 入力하여 컴퓨터로 하여금 標題上에 나타난 用語를 分類用파일에서 探索케 한 後, 最終적으로 이들 探索된 用語의 코드(code)를 一定의 分類 formula에 의하여 組合한 後 分類記號를 決定케 하는 것이다.

물론 標題上에 나타난 用語의 意味가 애매모호하거나 標題가 없는 경우에는 文獻의 內容을 分析하여 重要語(keyword)를 入力하여도 무방하나, 分類者에 따라 入力の 用語가 一致하지 않음에 따라 分類記號가 달라질 우려가 있으므로 可能한 이 方法은 피하는 것이 바람직하다고 하겠다.

2. 用語蒐集에 따른 코드화 方法

가. 用語의 蒐集

用語蒐集의 目的은 分類를 自動化하는데 있어서 컴퓨터로 하여금 文獻의 標題上에 나타나 있는 用語에 의한 探索으로 分類番號를 이끌어 낼 수 있도록 用語를 分析하고, 分析된 用語를 다시 코드화한 다음 分類用파일을 만들기 위한 것이다.

既存의 分類體系에 의한 分類는 모든 分類가 分類表를 근거로하여 이루어지나 自動化된 시스템의 分類는 이 分類用파일을 探索하여 分類番號가 分類原理에 따른 分類 formula에 의하여 決定되는 것이다. 따라서 用語의 蒐集은 既存의 分類方式에 있어서는 分類表가 必須적인 것과 마찬가지로 分類를 自動化하는데 있어서는 없어서는 안될 過程이라고 하겠다.

用語의 蒐集은 그 分野에서 使用되어 질 수 있는 大部分의 用語를 蒐集하여야 하며, 가능한한 同意語나 類似語등을 포함한 廣範圍한 蒐集이 바람직하다고 하겠다.

나. 用語分析

蒐集된 用語의 分析은 用語自體가 지니고 있는 屬性, 즉 그 概念을 分析하고, 分析된 概念下에서 一定한 記號를 부여함으로써 그 概念에 대한 代表置를 이끌어 내기 위함이다. 즉 蒐集된 用語는 分析하고 特定의 記號로 코

드화하여 컴퓨터의 자기디스크(magnetic disc)에 蓄積시켜 分類用과일을 維持함으로써 分類作業 時에 文獻의 標題를 入力하면 컴퓨터가 이 標題上에 있는 各用語를 分類用과일에서 探索하여 分類原理에 따른 組合의 過程을 거쳐 分類番號가 決定되도록 하기 위함이다.

用語의 分析은 同意語나 類似語가 아닌 경우에는 可能한한 固有番號를 부여하는 것이 바람직하다. 그 이유는 固有番號를 부여하면 固有의 分類番號가 形成될 수 있으므로 分類番號에 의한 配列뿐만 아니라 文獻의 自動檢索에 있어서도 正確性を 기할 수 있기 때문이다.

그러나 이렇게 固有番號를 전부 부여하게 되면 用語의 規模가 방대하여짐에 따라 分析된 用語에 대한 코드의 길이가 너무 길어지는 短點이 있다. 따라서 分析된 用語의 코드화는 가능한 한 個個用語마다 固有番號를 부여하여 分類番號에 의한 配列 및 文獻檢索의 側面을 考慮하여 同意語 및 類似語를 모아 같은 記號로 코드화하는 것이 바람직하다. 또 코드의 길이를 줄일 수 있는 方法으로서 十進數(0~9)에 의한 코드화만 適用할 것이 아니라 非十進數인 알파벳(a~z)을 함께 混用하여 作用하는 方法도 考慮할 수 있다고 하겠다.

蒐集된 用語의 分析方法은 다음과 같이 행할 수 있다.

① 두개의 名詞가 複合되어 하나의 概念을 形成한 경우에는 分離하지 않는다.

② 두개의 名詞가 複合되어 하나의 概念을 形成한 用語가 略語로도 表現하고 있는 경우에는 略語도 함께 使用한다.

③ 同意語 및 類似語는 같은 記號를 부여한다.

④ 形容詞+하나의 名詞로 이루어져 하나의 概念을 이루고 있는 경우에는 分離하지 않는다.

⑤ 形容詞+複合名詞의 경우에는 形容詞와 複合名詞로 分離하여 分析한다.

⑥ 세개의 名詞가 複合되어 하나의 概念을 表現하고 있는 경우에는 처음의 複合名詞와 다음의 한 名詞로 分離하여 分析한다.

⑦ 複數나 單數는 混用 使用한다.

⑧ 기타, 그 主題分野의 特殊性에 따른 各種 參照코드의 使用에 따른 分析을 수행한다.

다. 用語의 코드화方法

圖書館資料의 分類는 궁극적으로는 分類하지 않고 收入順으로 코드화하여 配架하는 方法도 考慮할 수 있으나, 이것보다는 自動化方法을 導入하여 自動적으로 關聯文獻을 grouping 시킴과 동시에 檢索에 있어 많은 接近點(access points)을 부여함으로써 效率的인 檢索이 되도록 하는 것이 앞으로의 圖書館自動化 추세에 부응하는 길이 될 것이다.

自動分類 時의 難點은 코드화인데, 어떻게 하면 分類作業 時에 人間의 人爲的인 조작이 없는 狀態에서 컴퓨터가 스스로 文獻의 標題만을 읽고 自動적으로 分類番號를 만들어 내느냐하는 것이며, 또 書架上에 配架하기 위하여 어떻게 간단한 記號를 만들어 내느냐 하는데 있다.

蒐集된 用語에 대한 코드화方法은 다음과 같이 十進記號式과 非十進記號式으로 番分하여 볼 수 있다.

1. 十進記號式

十進記號法은 0~9까지의 10個의 數단을 使用하여 코드화하여야 하기 때문에 主題分野의 用語數가 방대한 경우에는 코드의 길이가 길어진다.

例를 들어 한 主題分野에 있어 使用可能한 用語의 數가 總 4,500⁽³⁰⁾個이고, 이 가운데 [P]가 1,500個, [M]이 1,000個, [E]가 2,000個 등으로 되어 있다면 이에 따른 코드화가 可能한 用語의 總數(T)는 各 카테고리別 코드의 길이(n)에 따라 決定되며 이것은 $T=10^n$ 로 나타낼 수 있다. 이에 [P]카테고리의 用語는 코드의 길이가 4자리, [M]은 3자리, [E]는 4자리로 된다.

以上과 같이 十進記號式은 用語數에 따라 코드의 길이가 이어지는 短點이 있기 때문에 分類番號의 길이가 너무 길어지는 短點이 있다 하겠다.

2. 非十進式(混合式)

(30) 4,500個는 [S]와 [T]를 除外한 用語의 總數임.

22 도서관학논집

非十進式的 長點은 上記 十進式的 短點을 補充할 수 있다. 그러나 分類番號는 숫자(0~9)와 文字(a~z)로 構成되기 때문에 識別上에 있어 다소 어려움을 지니고 있는점이 短點이라고 하겠다.

코드화의 方法은 컴퓨터의 올림차순배열(ascending sort) 原則에 따라 숫자를 먼저 配列되도록 하고 文字를 다음 次順이 되도록 한다. 코드화의 예는 다음과 같다.

00	10	20	30	90	a0	b0	c0	y0	z0
01	11	21	31	91	a1	b1	c1	y1	z1
02	12	22	32	92	a2	b2	c2	y2	z2
}	}	}	}	}	}	}	}	}	}
09	19	29	39	~ 99	a9	b9	c9	~ y9	z9
0a	1a	2a	3a	9a	aa	ba	ca	ya	za
0b	1b	2b	3b	9b	ab	bb	cb	yb	zb
}	}	}	}	}	}	}	}	}	}
0z	1z	2z	3z	9z	az	bz	cz	yz	zz

上記와 같이 混合式으로 코드화할 경우 2자리 數로서 코드화할 수 있는 用語의 總數는 $36P1 \times 36P1 = 36 \times 36 = 1,296$ 이 된다. 따라서 各 카테고리別로 2자리로 코드화하면 總 $1,296 \times 5 = 6,480$ 個의 用語에 대하여 獨立的인 코드화가 可能하다는 結論을 얻게 된다.

十進記號式은 各 카테고리別 用語의 數가 1,000個를 넘으면 4자리, 1,000個까지는 3자로 表現하여야 하나, 이 混合式은 1,296個까지는 2자리로 表現可能하고 3자리로 表現하면 무려 46,656個의 用語에 대하여 固有番號를 부여할 수 있게 된다.

따라서 特定主題分野에 있어 分類用과일 作成을 위한 用語의 數가 各 카테고리別로 1,296(총 6,480個)個를 초과하지 않는다면 混合式의 경우에는 이들 個個의 用語에 대하여 2자리 數로 나타낼 수 있으며, 實際 코드화할 수 있는 用語數는 同意語나 類似語에 대하여 同一한 記號로 코드화하기 때문에 더욱 增加한다고 하겠다.

3. 分類用파일 設計

分類自動화를 위한 分類用파일의 設計는 크게 두가지 類型으로 區分할 수 있다. 첫째는 特定主題分野만을 위한 경우이고, 둘째는 KDC나 DDC와 같이 모든 學問分野의 主題를 모두 網羅할 수 있는 一般分類用파일의 設計이다. 이들 두 類型別 分類用파일의 設計方法을 보면 다음과 같다.

가. 特定主題分野의 分類用파일

여기에서 特定主題分野라고 하는 것은 CC 상에 나타나 있는 基本類(Basic Class)에 해당하는 것으로서 例를 들면 圖書館學, 農學, 心理學 등을 意味한다. 따라서 이러한 한 主題分野만을 위하여 分類自動화를 하려고 한다면 用語蒐集은 이 分野의 시소오러스(Thesaurus)를 參考하여 分析할 수 있으며, 란가나단이 個個學問分野마다 獨특하게 提示하고 있는 分類原理를 適用하여 分類할 수 있는 長點을 지닌다.

分類用파일을 위한 入力要素로서는⁽³¹⁾ ① 用語, ② 카테고리名, ③ 分析된 用語의 코드, ④ 參照코드, 등이 있다.

나. 一般分類用파일

一般分類用파일의 設計에는 다음의 2가지 側面에서 考慮할 수 있다. 하나는 文獻의 標題를 바로 컴퓨터에 入力하는 方法이고, 다른 하나는 入力時에 그 文獻에 대한 基本類를 標題와 함께 入力하는 方法이다. 前者의 方法은 完全自動化的 方法이기는 하나 지금 現在의 狀況으로서는 컴퓨터가 文獻의 標題만을 읽고 바로 그 文獻의 學問分野를 해독해 내기는 거의 不可能하다고 하겠다.

물론 自動分類가 文獻을 書架上에 配架할 記號를 自動적으로 부여해 내는데 있지만 같은 學問分野의 文獻을 인접시키지 못한다면 收入順 配架와 거의 다를바가 없기 때문에 分類自動化的 意味가 없어지게 된다. 따라서 가능

(31) 李慶浩, “圖書分類의 自動化—圖書館學 및 情報學分野 書誌分類를 中心으로”, 未刊行碩士學位論文, 慶北大學校 大學院, 1980, p. 66.

한 한 自動的으로 分類되도록 하되 같은 主題의 文獻을 인접시킬 수 있어야 하기 때문에 지금의 狀況에 있어 前者의 方法은 正確하게 論할 수가 없는 實情이라고 하겠다.

그리고 後者의 方法은 分類時에 基本類를 標題와 함께 入力하여야 하므로 약간의 手作業에 의하여 大略的인 分類가 행하여 진 後에 컴퓨터에 入力되는 경우라 하겠다. 그러나 이점은 앞으로 獻文을 出版할 당시에 그 文獻에 대한 基本類를 標題와 함께 記入하도록 함으로써 얼마든지 可能케 할 수 있다. 따라서 여기에서도 이 後者의 方法論에 입각한 一般分類用파일 作成에 대하여 言及하고자 한다.

特定主題分野를 위한 分類用파일은 個個 主題分野의 시소오러스를 對象으로 分析, 코드화 할 수 있지만, 一般分類用파일은 全學問分野를 網羅한 一般시소오러스가 存在하지 않기 때문에, 用語의 蒐集 및 分析은 個個 主題分野마다 이미 言及한 바 있는 特定主題分野의 分類用파일 作成과 同一한 方法으로 分析한 後 實現되어질 수 있다.

個個 主題分野마다 分類用파일을 위한 用語의 分析이 完成된 後에는 이들 個個 用語들에 대한 基本類와 코드를 함께 入力하여 알파벳 順으로 소오트(sort)한다. 소오트한 後 一般分類用파일을 만들게 된다 이 때 소오트한 後 分析된 用語가 만약 cooperation 이라면,

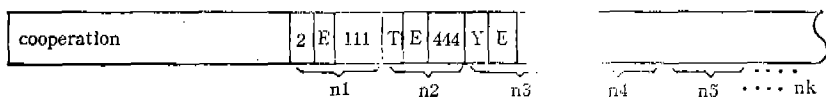
```

 cooperation 2 [E] 111
 cooperation T [E] 444
 cooperation Y [E] 112

```

등으로 나타나게 된다. 즉 cooperation 이란 用語는 圖書館學(2)에서는 [E] 111이고, 教育學(T)에서는 [E] 444이며, 社會學에서는 [E] 112로 分析, 코드화되어 있음을 意味한다.

따라서 이것을 근거로 하여 一般分類用파일의 設計포맷은 다음과 같이 設計할 수 있으며, 이에 따른 分類過程을 보면 圖書館學關係 資料를 分類할



때는 $n1$ 의 [E] 111을 適用하고 教育學의 경우에는 $n2$ 의 [E] 444를 適用하도록 하는 方法이다. 이 때 cooperation 이란 用語가 여러 學問分野에서 使用한다면 $n4, n5 \dots$ 등으로 더욱 늘어나게 된다.

4. 分類自動化的 基本原理

分類는 그 文獻에 대한 主題分析의 過程을 거쳐 그 文獻을 代表하는 用語를 抽出하여 이 用語에 대하여 適切な 記號를 부여함으로써 이루어지며, 이 때 主題分析은 주로 文獻의 標題나 內容 및 抄錄등을 分析對象으로 삼고 있다.

그러나 本 研究에서는 文獻을 代表한다고 할 수 있는 標題를 對象으로 하여 分類할 수 있는 分類自動化原理를 誘導하여 보고자 하였다. 이것은 무엇보다도 文獻의 內容이 標題上에 가장 最小의 用語로 가장 適切히 묘사되어 있을 것으로 確信하였기 때문이었다.

물론 文獻의 內容가운데에서 用語를 抽出하여 分類한다고 하여 分類되지 않는 것은 아니다. 이것은 어디까지나 入力의 內容을 各 圖書館마다 統一시켜 줄으로써 同一한 文獻에 있어서는 어떤 圖書館에서 分類하더라도 同一한 分類番號를 維持시켜 보고자 하는데 있을 뿐이며, 文獻의 標題가 그 文獻의 內容을 充分히 表現하고 있지 않을 경우에는 內容分析에 의해 抽出한 用語를 入力하여 分類하는 方法도 考慮할 수 있다 하겠다.

本 研究에서 誘導하여 낸 分類自動化原理는 크게 大部分의 學問分野에서 適用 可能한 一般共通原理와 特定分野에만 適用 可能한 應用原理의 두 部分으로 區分하여 볼 수 있으며, 이들에 대한 各 原理를 살펴보면 다음과 같다.

가. 一般共通原理

一般共通原理는 [P] [M] [E] [S] [T]가 다같이 使用되어질 수 있고, 단일

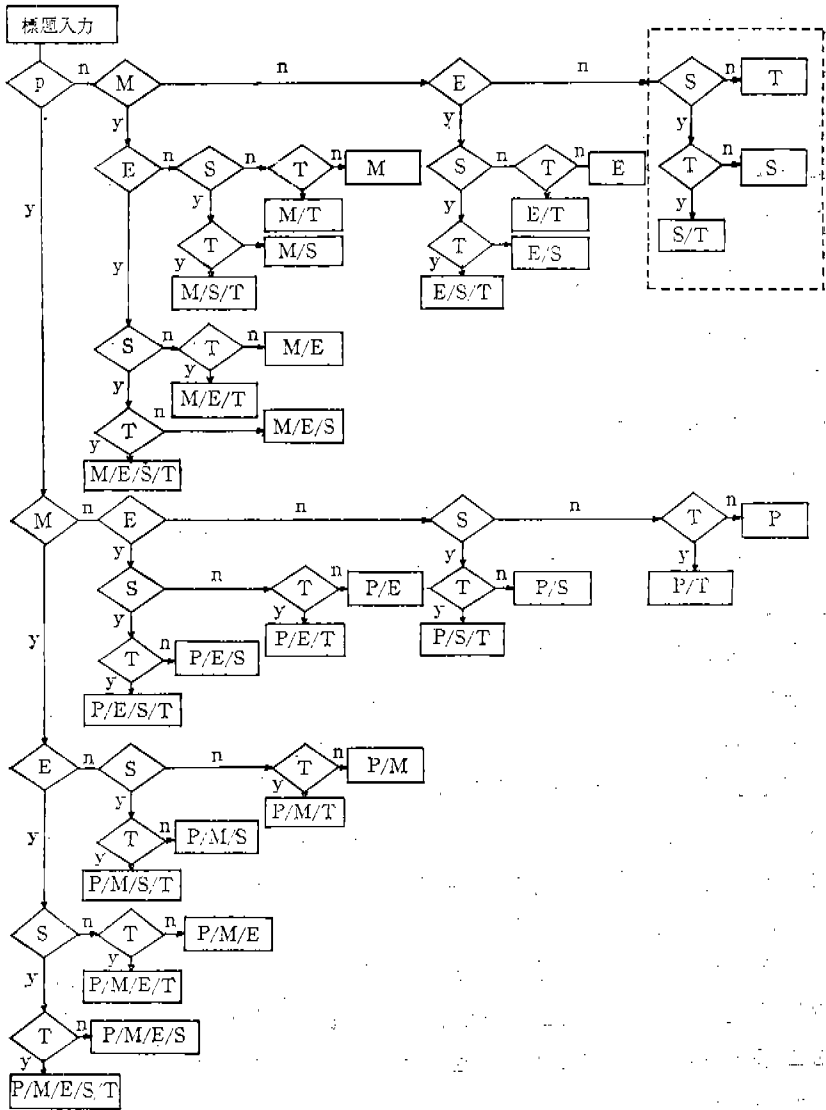


圖 3. 分類自動化的 一般共通原理

分類公式(formula)을 갖고 있는 分野를 意味한다. 따라서 이때의 分類原理는 다음 圖 3과 같이 나타낼 수 있다.

圖 3 上에 나타난 原理에 따르면 P.M.E.S.T의 5個 카테고리로서 組合되어 질 수 있는 경우의 수는 P/M/E/S/T, P/M/E/S, P/M/E/T, ...S/T, S, T에 이르기까지 모두 31가지 方法이 있으나, 각 카테고리마다 重複을 許한다면 그 경우의 수는 더욱 더 늘어나게 된다.

여기서 한 例를 들어 어떤 한 主題分野에서 [P]카테고리의 用語가 500個, [M]이 300個, [E]가 400個, [S]가 200個, [T]가 100個 정도 蒐集, 分析되어 分類用파일을 形成하고 있다면, 이들 用語에 의한 組合可能한 경우의 數는 다음 表 3과 같이 나타나 總 1, 227, 625, 851, 500이 된다. 이것은 1, 227, 625, 851, 500卷의 文獻에 대하여 分類番號의 重複이 없이 分類할 수 있는 수와도 같다고 하겠다.

그러나 [S]와 [T]는 사실상 그 自體만으로는 特定分野內에서의 基本카테고리가 아니기 때문에 標題가운데에서 [S]나 [T] 혹은 [S] [T]만으로 主題가 構成되는 例는 거의 없을 것으로 思料되어 진다. 그러나 [S]나 [T]의 組合으로 構成되는 경우라도 分類가 可能하게 된다.

以上과 같은 分類自動化의 一般共通原理에 따른 用語의 組合結果는 各 用語의 코드앞에 連結符號를 부여하여 分類番號가 決定되어 진다. 즉 [T]에는 '(apostrophe), [S]에는 .(dat), [E]에는 :(colon), [M]에는 ;(semicolon), 그리고 [P]에는 ,(comma)를 個個의 코드앞에 부여하되 [P]가 제일 처음 나타날 때에는 ,를 省略한다.

나. 應用原理

上記의 一般共通原理는 란가나단이 提示한 PMEST의 5個 카테고리에 의한 組合으로 이루어진 것으로서 一般的인 原理의 性格을 지니고 있다. 그러나 모든 學問分野에 있어서 반드시 이 原理만으로 分類되어지는 것은 아니다. 물론 이 原理에 의거하여 分類할 수 있을 지라도 이것은 分類의 基本目的인 資料의 類別化 하는데 있어서는 좋은 結果를 가져올 수 없다. 따라서

項 目	組合 formula	算 出 公 式	경 우 의 수
①	P/M/E/S/T	$500 \times 300 \times 400 \times 200 \times 100$	1,200,000,000,000
②	P/M/E/S	$500 \times 300 \times 400 \times 200$	12,000,000,000
③	P/M/E/T	$500 \times 300 \times 400 \times 100$	6,000,000,000
④	P/M/E	$500 \times 300 \times 400$	60,000,000
⑤	P/M/S/T	$500 \times 300 \times 200 \times 100$	3,000,000,000
⑥	P/M/S	$500 \times 300 \times 200$	30,000,000
⑦	P/M/T	$500 \times 300 \times 100$	15,000,000
⑧	P/M	500×300	150,000
⑨	P/E/S/T	$500 \times 400 \times 200 \times 100$	4,000,000,000
⑩	P/E/S	$500 \times 400 \times 200$	40,000,000
⑪	P/E/T	$500 \times 400 \times 100$	20,000,000
⑫	P/E	500×400	200,000
⑬	P/S/T	$500 \times 200 \times 100$	10,000,000
⑭	P/S	500×200	100,000
⑮	P/T	500×100	50,000
⑯	P	500	500
⑰	M/E/S/T	$300 \times 400 \times 200 \times 100$	2,400,000,000
⑱	M/E/S	$300 \times 400 \times 200$	24,000,000
⑲	M/E/T	$300 \times 400 \times 100$	12,000,000
㉑	M/E	$300 \times 400 \times 120 \times 400$	120,000
㉒	M/S/T	$300 \times 200 \times 100$	6,000,000
㉓	M/S	300×200	60,000
㉔	M/T	300×100	30,000
㉕	M	300	300
㉖	E/S/T	$400 \times 200 \times 100$	8,000,000
㉗	E/S	400×200	80,000
㉘	E/T	400×100	40,000
㉙	E	400	400
㉚	S/T	200×100	20,000
㉛	S	200	200
㉜	T	100	100
		계	1,227,625,851,500

特殊한 主題分野는 上記의 一般共通原理에 의거하여 分類할 것이 아니라 그 學問分野의 特性에 맞는 組合原理를 誘導해 내어야 할것으로 思料된다.

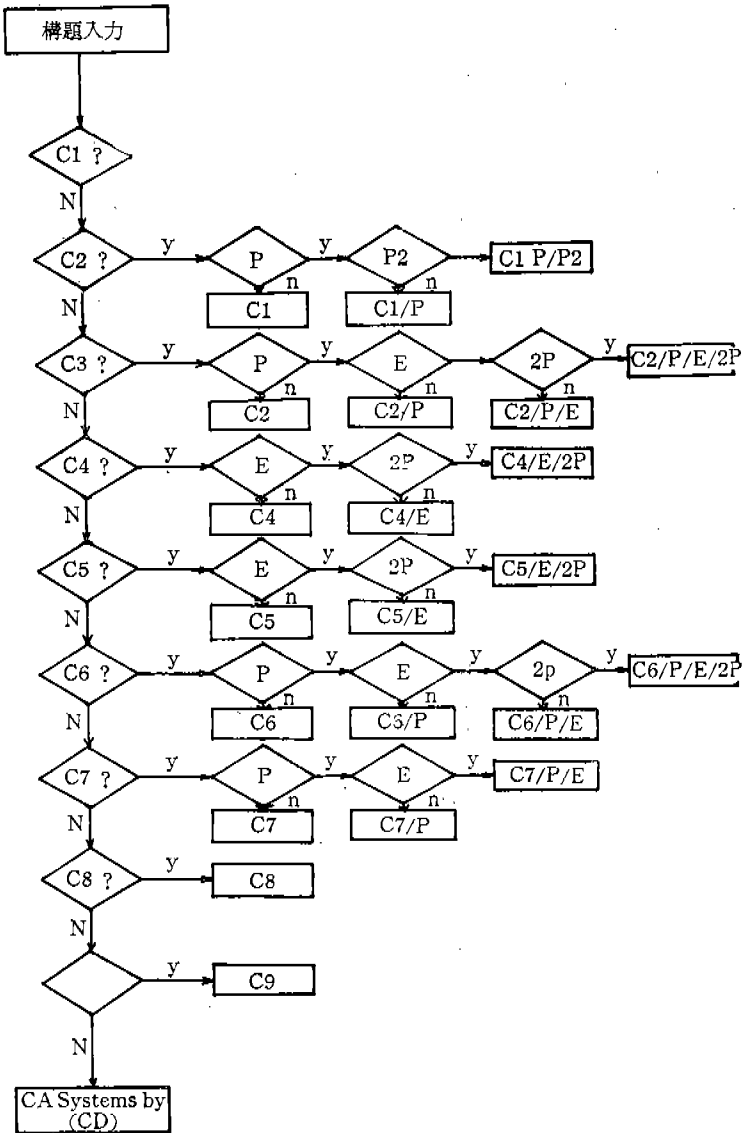


圖 4. 物理學分野 分類自動化 原理

이와같이 特殊한 原理를 性要로 하는 學問分野는 上記 表 2 上에 나타난
 마와같이 分類公式(formula)이 [P] [M] [E] [S] [T]로 構成되어 있지 않고
 [P]카테고리만으로 構成되어 있는 경우나, [M]카테고리가 없이 分類되어
 지는 경우, 또 한 主題分野內에서 여러 개의 分類公式를 適用하고 있는 경
 우 等이다.

따라서 이 應用原理는 위와 같은 內容을 包含하고 있는 學問分野마다 各
 各 그 原理를 誘導하여 내어야 하나 本 研究에서는 物理學分野를 이의 한
 例로서 說明하고자 한다.

物理學分野를 CC의 分類表上에서 보면 Canonical Division이 C1 Fun-
 damentals, C2 物質의 性屬, C3 소리(Sound), C4 열(Heat), C5 빛, 방사
 선(Light, Radiation), C6 전기(Electricity), C7 磁氣作用(Magnetism),
 및 C8 Cosmic hpyothese로 區分되어 있다. 그리고 이에 따른 分類公式이
 C2는 C2[P], [P2], C3는 C3[P]:[E] [2P], C4는 C4:[E] [2P], C5는 C5
 [P]:[E] [2P], C6는 C6[P] [E] [2P] 및 C7는 C7[P]:[E]로 形成되어 있다.

따라서 이 物理學分野에서의 分類公式은 物理學이란 독특한 學問의 特性
 으로 인하여 物理學의 個個 分野마다 各各 다른 分類公式를 適用하고 있다.
 物理學分野의 分類自動化 原理를 誘導하여 보면 다음 圖 4와 같다.

V. 適用上의 問題點 및 提言

1. 標題의 表現

文獻의 標題는 그 文獻이 무엇에 관하여 記錄하고 있는지를 가장 간단 명
 료하게 表現하여야 하기 때문에 단지 指示의 情報만으로 나타낸다. 물론 文
 獻의 標題은가에는 애매모호한 것도 있지만 大部分은 文獻의 內容을 充分
 히 나타내고 있다고 하겠다.

특히 標題入力에 의한 分類에 있어서는 檢索의 機能도 함께 지니고 있으
 므로 너무 包括的인 用語도 피하여야 하겠지만, 너무 많은 用語에 의한 表

現도 分類番號의 길이 가 길어지는 短點이 있으므로 考慮對象이 된다고 하겠다.

따라서 이러한 점을 考慮하여 文獻의 內容을 가장 適切히 나타낼 수 있는 具體的인 用語로 標題가 記述되어 질 수 있도록 著者나 出版關係者에게 그 必要性을 널리 周知시키는 것이 바람직하다고 하겠다.

2. 主題識別코드의 부여

標題入力에 의한 完全한 分類自動化가 이루어지기 위해서는 個個 文獻에 대하여 主題識別코드 혹은 用語를 부여하여야 한다. 물론 主題識別코드나 用語없이도 分類할 수 있지만 가능한 한 같은 主題의 文獻을 한곳에 모으기 (grouping) 위해서는 個個 文獻의 標題와 함께 關聯學問分野를 明示하면 보다 正確히 分類할 수 있다. 따라서 完全한 分類自動化를 위해서는 文獻出版 時에 主題識別코드 혹은 用語의 부여가 要望된다고 하겠다.

3. 各 學問分野의 分類自動化原理定立

分類自動化는 文獻의 標題를 바로 入力하여 컴퓨터로 하여금 이 標題가운데에서 主題(基本類)를 認識함으로써 分類가 始作되며, 이 때 分類過程은 그 學問分野의 分類公式에 따라 행하여 진다.

따라서 一般共通原理를 適用할 수 없는 特殊한 學問分野는 그 學問分野의 特殊性에 따라 特別히 分類自動化의 原理(應用原理)를 定立하여야 할 必要性이 있다고 하겠다. 이의 한 例가 바로 上記의 物理學分野 分類自動化 原理이다.

4. 分類番號의 길이

分類番號의 길이는 同意語나 類似語 및 各種 參照코드에 의한 組合이 없는 한 入力되는 用語의 數에 따라 비례하기 때문에 標題가 긴 경우에는 그만큼 分類番號의 길이가 길어진다. 이 分類番號의 길이를 줄이는 가장 最善의

方法은 非十進記號의 使用인데, 이것은 書架上的 配架 時에 많은 不便이 초래되는 短點이 있다. 따라서 檢索과 配架가 다같이 용이하고 分類番號의 길이도 단축시킬 수 있는 코드法의 開發이 要望된다고 하겠다.

5. 團體名 및 人名

分類用파일 作成 時에 團體名이나 人名까지 包含하기란 거의 不可能하다. 또 團體名의 경우에는 대체로 그 表現自體가 긴 文章으로 되어 있기 때문에 概念의 組合으로 分類番號를 만들어 낼 수 없으며, 人名도 個個의 코드를 부여할 수 없다고 하겠다.

따라서 分類가 完全히 自動化 되려면 이러한 點도 考慮하여야 할 것으로 思料되어 진다.

Ⅶ. 結 論

圖書館業務 가운데 分類를 除外한 거의 모든 業務는 점차 自動化되어 가고 있다. 그러나 分類만은 아직까지도 自動化를 위한 理論的 基盤이 구축되지 않고 있으며, 어떤 圖書館도 完全히 自動化된 圖書館은 찾아볼 수가 없는 實情이다.

이에 本 研究는 이와같은 當面問題를 다소나마 解決할 수 있는 실마리를 提供하여 보고자 하는 意圖에서, 分類自動化를 위한 原理誘導에 관한 하나의 方法論을 CC의 原理에 입각하여 提示하여 보았다.

本 研究의 結果를 要約하면 다음과 같다.

1. 本 研究는 分類自動化를 위한 原理誘導에 관한 理論的 研究이다.
2. 分類自動化를 위한 原理誘導는 CC의 原理에 입각하였다.
3. CC의 原理에 입각한 分類自動化의 原理는 一般共通原理와 應用原理의 兩部分으로 誘導하였다.
4. 分類自動化를 위한 分類用파일의 設計는 特定主題分野를 위한 分類用

파일과 一般分類用파일로 區分하여 設計하였다.

5. 分類의 方法은 文獻의 標題를 컴퓨터에 바로 入力함으로써 分類되어질 수 있도록 하였다.

또 本 研究에서 誘導한 分類自動化 原理에 의해 分類自動化가 實現될 경우 다음과 같은 效果를 期待할 수 있다.

1. 分類의 迅速, 正確性을 기할 수 있다.
2. 어떤 圖書館에서 分類하더라도 같은 文獻은 하나의 共通된 分類番號를 갖게 한다.
3. 迅速한 分類로 利用者에게 最大限의 奉仕를 할 수 있다.
4. 利用者는 자기가 원하는 主題의 資料를 一括 檢索할 수 있다.
5. 概念組合 方法이므로 複雜한 主題의 表現도 容易하다.
6. 圖書館의 完全 自動化에 기여할 수 있다.

參 考 文 獻

註로 代身함.

Principles of the Automatic Book-Classification

Shim, Eui Soon* Lee, Kyung Ho*

〈Abstract〉

The purpose of this study is to build a general principle for the automatic book-classification which can be put to use in library operation, and to present a methodology of the automatic classification for the library.

Since the enumerative classification scheme which exist as manual systems cannot be applied to the automation of classification, the principles of Colon Classification by S.R. Ranganathan is brought in and studied.

The result of the study can be summarized as follows:

- (1) Automatic book-classification can be performed by the principles of faceted classification.
- (2) This study presents a general and an application principles for the automatic book-classification.
- (3) File design for the automatic book-classification of a general classification is different from that of special classification,
- (4) The methodology is to classify the literature by inputing the title into a terminal.

In addition, the expected Value from the Automatic Book-Classification is as follows:

- (1) The prompt and accurate process of classification is possible.

* Dept. of Library Science, College of Social Science, Teagu University

(2) Though a book is classified in any library it can have the same classification number.

(3) The user can retrieve the classification code of a book for which he or she wants to search through the dialogue with the computer.

(4) Since the concept coordination method is employed, even the representing of a multi-subject concept is made simple.

(5) By performing automatic book-classification, the automation of library operation can be completed.