

◎ Information Science in Librarianship; with Emphasis on Some Problems in the Information Processes in the Library

圖書館의 情報奉仕

—그 準備過程에서 일어나는 몇가지 問題에 대하여—

崔 成 眞

(成均館大學校 圖書館學科 教授)

<차 례>

1. 緒 言
2. 情報시스템이란?
3. 圖書館과 情報센터
4. 情報의 處理過程
 - 4.1. 데이터 入力過程에서 일어나는 問題
 - 4.2. 蓄積 및 檢索過程에서 일어나는 問題
 - 4.3. 응답자료 提供過程에서 일어나는 問題
5. 結 言

1. 緒 言

社會가 복잡해짐에 따라 意思決定의 過程이 복잡해지고 이것이 다시 情報奉仕의 必要를 增大시켜왔음은 周知의 사실이다. 비록 역사는 짧지만 컴퓨터라는 것이 그 동안, 방대한 자료를 高速으로 처리할 수 있는 能力을 상당히 開發한 것도 잘 아는 사실이다. 이 두가지 발전 즉 增大된 報情需要와 增大된 情報處理能力이 結合하여 지난 四半世紀 동안 各種 情報시스템이 많이 생겨났다. 科學技術界에서 「情報센터」라 부르고 社會科學分野에서 「데이터·뱅크」 企業部門에서 「管理情報센터」 軍隊에서 「指揮統制센터」라 하는 것이 다 그것들이다. 여러 情報시스템이, 그 시스템設計나 意思決定過程과의 有機的 連結等에 아직도 問題가 남아 있기는 하나 다소 부족함대로 각각 實用단계에 들어선지 오래이다. 檢討와 改善이 進行되고 있으므로 問題들은 早晚間 해결될 것이고 여러가지 種類의 情報시스템이 가까운 將來에 더 많이 생겨날 것이 豫想되고 있다.

이러한 情報시스템들이 오늘 날 傳統的 圖書館 방식에 대하여 도전을 가해 오고 있다. 그리하여 아슈르바니팔 王朝以來 司書들의 聖域이던 圖書館은 처음으로 危機를 맞아 동요를 일으키고 있다.

이러한 때 圖書館人들이 할 일이 무엇이겠는가를 한번 생각해 볼 필요가 있다. 무엇보다, 情報시스템이란

무엇이며, 왜 그런 것들이 나타나게 되었으며, 社會에 어떤 貢獻을 할 수 있는지를 바르게 認識하고 도전에 응할 태세를 갖추는 일이 時急하지 않을까?

유감스러운 일이지만 圖書館의 使命과 機能에 對해서는 아직도 萬人共感하는 單一概念의 提示가 없다. 圖書館이란 단순히 圖書를 保存하는 곳인가? 아니면 情報奉仕를 하는 곳인가? 媒體와 內容 중 그 어느 것이 圖書館에서 더 重要한가? 이런 質問들은 사실 심심풀이 話題 이상의 심각한 의미를 내포하는 것인줄 안다.

이 質問들이 처음 공식적으로 提起된것은 1847年에서 1849년까지 繼續된 大英博物館圖書館委員會의 「著者·書名目錄이나 主題目錄이나를 決定하기 위한 公聽會」에서 였다¹⁾ 當時 圖書館의 使命은 過去記錄을 保存하여 未來에 전해주는 것이 全部이었으므로, 두가지 目錄이 다 必要한것이기기는 했지만 擇一해야 하는 形편에서 著者·書名 目錄으로 決定했던 것이다. 어떤 特定圖書의 所藏與否를 밝히는 데는 著者·書名 目錄이 有用하고 또 그것이 當時의 「圖書館」性格에 여러모로 맞았으므로 이를 만들기로 했던 것이다. 그때 著者·書名 目錄을 備置하기로 한 사실이 그후 한 세기 이상 英美 圖書館과 世界 여러 圖書館의 性格을 決定하게 되리라고는 決定參與者들도 아마 몰랐을는지 모른다.

날로 增大되는 報情需要는 이제 圖書館으로 하여금 百餘年前에 한 決定을 再評價하도록 압력을 가하고 있다. 分明히 過去記錄保存者로서의 圖書館의 役割은 오늘도 變함없이 重要하다. 情報란 抽象 속에 存在하는 것이 아니고 어떤 형태로든 記錄物로 存在하는 것이므로 情報만을 원하는 사람에게도 그것이 記錄된 文獻

1) Great Britain, Commissioners Appointed to Inquire into the Constitution and Government of the British Museum, Report. London, HMDS, 1850 ("The Pan-nizi Report").

이 있어야 하는 것이다. 그렇다고 圖書館이 계속하여 文獻保存을 가장 主要한 機能으로 간주해야 할 것인지?

이에 대한 답은, 可否間, 앞으로의 圖書館性格이나 奉仕스타일에 또 한번 큰 영향을 미치게 될 것이 아닌가 생각된다. 重大한 決定의 時期에 처하여 圖書館人들은 도전자와 正體 即 情報의 性質, 行動, 流通現象 등을 可及의 많이 배우고 많이 알아야 할 것이라 믿어진다. 圖書館 테두리 안에 導入할 수 있는 情報奉仕의 種類 範圍 方法等도 검토해야 할 것이다.

그 동안 이에 대하여 圖書館界에서 서서히 注意를 기울여 온 것은 사실이다. 圖書館學科 教育內容이 急激히 變化되어가며 새롭고 異質의 情報學科目들이 導入되고 있다는 사실만 하더라도 圖書館人들이 時代的 要求에 응답하여 움직이고 있다는 증거라고 할 수 있다. 情報學은 이제 圖書館學教育의 한 重要한 部分이 되었으며 司書職에서 뱉 수 없는 教養으로 간주되고 있다.

情報學에 대하여 이렇게 圖書館學界에서 理論的으로 前進의 受容態勢를 보이고 있음에도 불구하고 圖書館의 실제 봉사활동면에서는 情報奉仕를 위한 준비가 매우 지지부진한 실정이다. 앞에서도 若干 言及한 바와 같이 歷史的으로 보면 圖書館은 資料——主로 圖書——를 物體로 保存하여 再利用하는 것이 主使命이었다.

그러한 業務는 必要에 대하여 受動的이고 서어비스 形態는 靜的인 것이다. 人文·社會科學 爲主의 前世紀까지는 그런 狀態로 圖書館에 대한 社會的 期待가 대충 充足되었다고 볼 수 있다. 그러나 時代와 더불어 圖書館利用者의 必要와 周圍環境은 變하는 것이다. 오늘 날 科學技術 分野의 擴大와 專門化, 研究開發部門의 競爭等은 圖書館利用者의 性格과 必要를 크게 變化시켜 놓았다. 利用者의 必要가 서어비스·스타일을 決定한다 함은 現代圖書館奉仕原則의 第一條인 것이다. 圖書館은 여러가지 情報處理技術을 開發하여 利用者 必要에 對해 가장 適合한 方法으로 奉仕해야 할 것이다.

傳統的 圖書館方法이 最適일 경우도 있을 것이고 圖書資料보다 情報 自體를 要求하여 參考奉仕 書誌提供, 抄錄 索引提供, SDI 등의 형태가 알맞는 경우도 있을 것이다.

이러한 意味에서, 情報란 무엇이며 또 情報學이란 무엇인가, 그리고 圖書館이라는 既存 시스템 안에 情報奉仕를 導入하는 과정에 에견할 수 있는 主要問題는 무엇이었는가 등을 한번 생각해 볼 必要가 있을 것 같다.

2. 情報시스템이란?

情報시스템이 무엇인가를 말하기 전에 먼저 「情報」의 定義부터 생각하는 것이 순서일 것이다. 情報가 여러가지 뜻으로 쓰임은 周知의 사실이다. 投術者와 哲學者가 가지는 「情報」의 概念이 다른지 모른다. 어떤 사람들은 情報란 記錄된 사실이라고 한다. 著述의 內容 또는 사람의 記憶 속에 들어간 경험이라고 하는 사람들도 있다.

이런 혼돈 속에서 하나 더 내 놓으며 情報의 定義云云하는 것은 최악인줄이나 情報시스템이 情報를 다루는 이상 基本的 水準에서라도 定義問題를 안 생각할 수 없다.

먼저 데이터, 정보, 지식, 지혜 등 몇가지 關聯用語들을 생각하고 이들 간의 關係를 考察하는 것이 情報의 定義를 유도하는 한 方法이 되지 않을까 생각한다.

데이터(data)는 일반적으로 情報를 만드는 原料라고 생각한다. 데이터는 반드시, 어떤 형태로 기록되어 처리할 수 있으며 事實(facts)을 나타낸다.

그러므로 이를 거꾸로 말하면 (뒤에 다시 구체적으로 정의하겠지만) 情報란 데이터를 처리한 결과인 것이다. 한편 知識이란 累積되고 體系화된 데이터이고 智慧는 知識과 區別되는 것으로 主觀的 倫理的, 判斷的인 것이다. 知識이 科學의 領域인데 비해 智慧는 人文 領域이라 할 수 있을 것이다. 이러한 關係를 單純化하여 다음과 같은 便宜上의 定義를 지어본다.

정보란 데이터를 처리한 結果로 얻는 데이터이다. 여기 「처리」한다는 것은 단순히 데이터를 전달한다는 뜻일 수도 있고, 데이터의 選擇, 데이터의 組織, 데이터의 分析이라는 뜻일 수도 있다.

정보도 데이터와 같이 記錄되므로 물론 다시 처리될 수 있다. 그러므로 情報와 데이터의 差異는 絕對的인 것이 아니고 同一한 자료라도 利用者에 대한 有用성에 따라 그것이 데이터가 될 수 있고 情報가 될 수 있는 그런 相對的인 것이다.

이 定義에서 한 가지 重要한 사실은 情報는 마치 生命과 같이 그것을 維持하는 部分이나 過程과 떨어져서 存在할 수 없다는 것이다. 이 사실은 情報는 반드시 어떤 特定한 「情報시스템」 테두리 안에서만 研究될 수 있음을 뜻하는 것이다.

이제 情報시스템에 대해 좀 생각해 본다. 먼저 어떤 사람이 複雜한 現象(시스템) 하나를 研究하기로 作定

2) Wiener, Norbert, *Cybernetics, or Control and Communication in the Animal and the Machine*, New York, Wiley, 1961.

하였다고 가정한다. 아마도 그는 그 현상을 分析하여 몇가지 面(aspects), 예컨대 그 外部構造라든지 사이버네틱스 經路³⁾ 新陳代謝過程, 情報處理過程 등으로 나누는 다음 그 하나 하나를 觀察 測定하여 밝혀 나가려고 할 것이다. 사람을 한 시스템으로 보고 그런 측면으로 研究하기로 하였다면 역시 한 사람을 자기 다른 機能을 하는 몇가지 面으로 나누어 이해하려고 할 것이다. 즉 骨格과 筋肉의 複合體로서, 밥과 물과 空氣를 處理하여 에너지로 變化시키는 化學工程으로서 그리고 수시로 미터를 받아드리고 決定을 지으며 人體의 여러 面을 統制하는 思考하는 사람으로서 分析할 것이다. 圖書館은 外部構造(施設, 資料, 資料의 흐름), 行政組織, 情報處理機關(資料의 수집, 정리, 대출)의 세 面으로 分析될 수 있을 것이다. 시스템에 대한 연구는 각 시스템에 適合한 研究方法를 擇하여 하게 된다. 生命體를 研究함에 生物學을 利用하고 行政組織을 研究함에 當然히 組織管理理論을 利用하는 것이다. 研究結果는 各種 모델이나 原理, 法則 등으로 나타나고 다시 이것들을 통해 시스템의 현상이 說明되고 未來의 현상이 豫見된다. 한 시스템 안에서 기능을 달리하는 面들 가운데 情報處理面이라고 부를만한 것들이 있는 수가 있다. 이 情報處理面이 餘他 面들보다 중요한 것일 때 그 시스템은 다른 아닌 情報시스템인 것이다.

그러므로 情報시스템이란 한 시스템(自然現象이나 어떤 構造라도 좋다) 안에 있는 情報를 處理하고 情報를 만들어 내는 面들을 總稱하는 말이라고 할 수 있다.

情報시스템의 一般型은 도표 1과 같다³⁾. 도표 1에 나타난 바와 같이 情報處理 및 生産過程은 다음과 같은 4단 單位로 나눌 수 있다. 즉 입력과 출력을 위한 커뮤니케이션, 情報의 蓄積, 情報의 處理, 피이드백의 4단 單位이다.

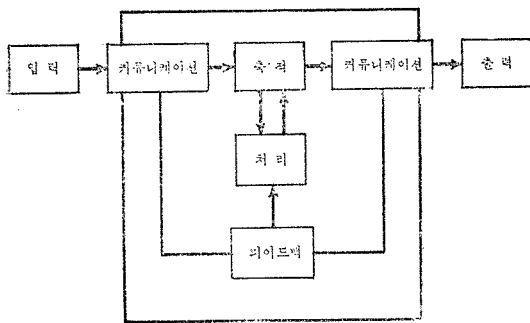


도표 1. 정보시스템

3. 圖書館과 情報센터

정보는 그것이 生産되어 讀者에 이르기까지 대체로 다음 네 단계를 거치게 된다.

- (1) 情報의 生産——研究者의 領域
- (2) 情報의 配布——情報生産者 또는 저어날 및 圖書 편집자의 領域
- (3) 情報의 選擇, 收集, 索引, 抄錄 및 目錄——司書, 文獻士 등의 領域
- (4) 情報의 分析——研究者 또는 情報專門家の 領域

위의 內容을 자세히 검토해 보면 圖書館이 情報活動過程에서 매우 重要的 役割을 맡고 있음을 알 수 있다. 資料의 選擇 收集者이며, 目錄 索引의 作成者이며, 그 파일의 蓄積 保存 奉仕者인 圖書館은 위의 (3) 단계를 사실상 專擔한 실정이다. 過去記錄의 保存者로서 圖書館은 먼저 어떤 자료를 保存할 것이냐를 결정하게 되므로 情報의 選擇을 담당하게 되고, 또 圖書館이 著者·書名目錄을 作成하면 그것이 情報“傳達”의 한 방법이므로 곧 初步의 이기는 하나 所藏資料에 대한 情報“處理”를 하는 것이며, 主題目錄을 作成하게 되면 이는 所藏資料 속에 있는 情報 選擇을 함을 의미하는 것이다.

그러나 圖書館目錄은 아무리 完全하게 만들어진 것이라도 資料內容의 대강 구분밖에 提示할 수 없다는 限界가 있다. 一般公共圖書館利用者들에게는 이것으로도 불평이 없을지 모르지만 同一한 主題分類 방식을 高度의 專門文獻 정리에 적용한다면 그런 일이 헛된 수고에 지나지 않을 것은 明白한 일이다. 主題를 細分할 수 있는 索引조차도 잠재적 관계문헌을 指示해 줄 뿐 內容을 알려주지는 못한다. 그러므로 책이나 저어날 記事와 같은 資料全體가 아니라 그 속에 있는 斷片的인 미터를 選別할 수 있는 새로운 方法을 開發하기 위해 그 동안 많은 노력이 傾注되어 왔다.

藏書 中에서 資料(document)를 찾아 내어 提供하는 시스템과 高度의 情報處理를 하는 시스템의 差異를 最初로 力說한 사람은 바아·힐필이었다. 그는 자료를 蓄積하고 檢索하는 문제는 마땅히 그 속에 포함된 미터를 分析하는 問題와는 別途로 研究되어야 한다고 主張하였다⁴⁾. 文獻檢索이란 어떤 選定된 主題에 關係 있는 資料가 어느것인지를 決定하는 행위에 지나지 않으나 情報檢索이란 選定된 主題에 關한 質問에 답을 얻어내는 행위라는 것이 그의 結論이다.

3) Broko, Harold, *The Analysis and Design of Information Systems*, Los Angeles, System Development Corporation, 1966.

4) Bar-Hillel, Yehoshua, *Language and Information: Selected Essays on Their Theory and Application*, Reading, Mass., Addison-Wesley, 1964.

情報處理시스템의 複雜度가 더해감에 따라 해당 연구 분야의 實質的 知識이 점점 더 깊어지게 되었다. 이런 需要에 따라 나타난것이 한 두 專門分野에 精通한 情報處理者인 소위 情報專門家(information specialist)이다. 情報專門家は 文獻속에 발표되는 研究結果를 分析하고, 새로운 研究結果들을 綜合評價하여 기술현황보고 記事(state-of-the-art report)를 쓰고, 未出版 文獻들을 수집하여 組織하는 등의 일을 한다. 그리고 이들이 모여 일하는 情報센터라는 새로운 기관이 出現하였고 그 活動內容은 날로 擴大되어가고 있다. 高度의 專門分野 知識을 通하여 情報센터나 그 밖의 유사한 情報시스템들은 圖書館이 하기 어려운 水準의 情報處理 및 奉仕를 담당하고 나선것이다.

4. 情報의 處理過程

典型的인 二次資料의 蓄積 및 檢索시스템을 생각해 본다. 그 處理過程은 도표 2와 같을 것이다.

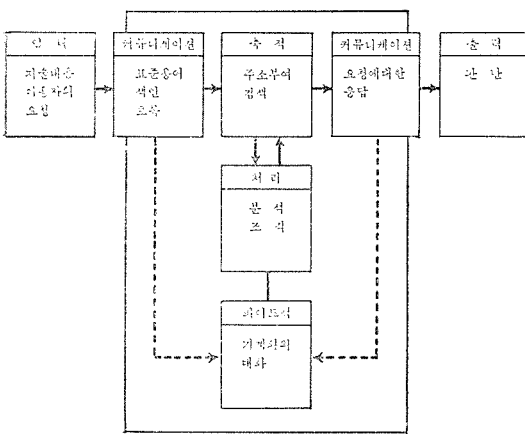


도표 2. 이차자료 축적 검색시스템

이 시스템의 第一段階인 커뮤니케이션 段階에서는 수집 文獻과 利用者要請의 入力을 위해 文獻과 要請이 記述目錄, 主題目錄, 索引 抄錄等 그 시스템에서 採擇한 「言語」로 表示되어야 한다. 이 入力原形 준비는 特別히 留意할 必要가 있다. 萬一 여기 잘못이 생기면 그 시스템 言語와 바르게 작어지지 못하고 동떨어지게 되므로 그 후의 모든 處理過程에서 아무리 完璧을 期한다하더라도 그것은 다만 헛된 수고가 될 뿐이다. 情報시스템 研究에 있어 入力과 시스템 사이의 대화(dialog)가 언제나 매우 重要한 分野로 간주되는 것은 바로 이때문인 것이다. 대화의 잘 잘못은 곧 시스템의 完 不完全을 決定하는 함수관계라 할 수 있다.

第二段階인 蓄積段階에서는 既存索引이나 抄錄파일 에 수집 文獻의 새 住所가 정해진다. 이 住所는 이후, 利用者의 要請에 따라 찾아 갈 住所이기도 하다. 適合한 住所를 찾아서 붙였는지 與否가 그 시스템의 効率度를 決定한다함은 說明을 要하지 않을 것이다.

住所를 정해준다는 것은 既存 파일속에 한 位置를 定義해준다는 이상의 아무것도 아니다. 이것은 새로 入力되는 文獻의 경우에 거치는 과정이지만 利用者가 어떤 要請을 提出했을 경우 파일 속의 既存 住所로 迅速 正確하게 찾아가서 記錄內容을 읽는 過程이 또 있다. 이 때의 「迅速」「正確」(그리고 理想的이려면) 「적은 費用으로」라는 條件들은 자연 여러가지 문제들을 提起한다.

住所를 찾아 가 읽은 記錄의 內容은 要請이나 文獻과의 合致度가 評價되어야 한다. 이 合致度 해석 상의 論理的 問題도 결코 작은 것이 아니다.

前述한 書誌記述, 資料位置의 探索, 文獻評價等은 다 圖書館에서 오랫동안 해 오는 일들이다. 그러나 그 다음 段階인 데이터 分析 및 要約過程은 이런 種類의 情報시스템에 附加되는 경우가 흔하지 않다. 아마도 分析 또는 要約 過程이 매우 複雜하기 때문일 것이다. 그 다음 段階는 所藏파일에 基礎한 積極的 周知活動이다.

初歩的인 二次資料 提供業務는 比較的 쉬운 일이며 실제로 여러 圖書館에서 實施되고 있다. 가장 흔한 例로 論文標題索引 같은 것을 들 수 있을 것이다.

위에서 말한 過程들이 圖書館分野 안에서의 情報學의 主要 關心事들이 아닌가 생각한다. 아래에 圖書館의 情報奉仕를 研究해 온 사람들이 指摘한 몇가지 問題들을 中心으로 위의 各 段階에 提起되는 理論的인 또는 實際的인 問題들을 要約해 보고자 한다. 勿論 그 많은 問題 하나 하나의 해결까지 여기서 생각할 수는 없다. 또 모든 問題를 망라적으로 다룰 수도 깊이 다룰 수도 없다. 다만 이 分野에 關心을 가지고 있는 圖書館人들에게 文獻을 통해 더 깊이 연구할 수 있는 方向 提示나 되었으면 할 뿐이다.

4.1. 데이터 入力過程에서 일어나는 問題

제일 처음으로 부딪히는 문제는 데이터 入力을 위한 커뮤니케이션 過程에서 일어나는 것으로 그 중의 하나는 도서관의 主題標目典據作成問題에 해당하는 用語標準化問題이다. 이 問題는 具體的으로 辭典, 디스오러스, 分類表 等を 作成하는 문제이다. 또 한 문제는 機械와의 對話問題로 圖書館에 이에 해당하는 문제는 없으나 서로 말이 통하지 않는 어떤 두 사람의 對話問題와도 같은 것이라 생각할 수 있다. 세번째는 索引, 抄

錄, 記號化(coding) 문제인데 이는 編目에 「해당」한다. 「해당」한다고 하지만 실상 이들과의 차이는 細分의 程度 問題이고 本質的으로 兩者間에 다른 것은 없는 것이다.

用語標準化 問題

어떤 情報시스템이든지(機械化된 것이건 안된 것이건 간에) 그 基本을 形成하는 것은 標準用語로 記錄된 데이터의 蓄積이다. 여기 한 말로 「標準用語로」라고 하였지만 그 標準化된 程度는 勿論 各各 다를 것이다.

用語의 制定 또는 選定問題는 오랫동안 言語學者, 哲學者 心理學者 等の 共同關心事였다. 이들은 各己 다른 角度에서 情報시스템 用 標準用語를 생각하였다.

言語學者들은 觀察을 통해 이 문제를 해결하려고 試圖했다. 즉 어떤 特定 情報시스템에 속한 사람들은 어떤 말을 쓰는가를 알아내려고 한 것이다⁵⁾.

한편 哲學者들은 言語의 標準化를 論理的인 問題로 보았다. 즉 言語의 基礎가 되어 있는 構造를 알아냄으로 標準化問題는 해결된다고 생각한 것이다⁶⁾.

心理學者들은 若干 다른 方向을 취하였다. 言語가, 말하는 本人과 그것을 듣는 相對方 사람에게 미치는 영향이 각각 어떤 것인가에 關心을 集中한 것이다.

情報專門家들은 言語의 標準化에 關心을 기울여 왔다. 이들은 주로 컴퓨터와의 커뮤니케이션 過程에 쓰일 標準言語를 開發하려고 한 것이다.

研究方向은 어떻게 잡았든지 그 目的은 다 모호한 相互關係의 말들을 윤곽이 分明한 형태로 傳達하려는 것이다. 그런 形式化된 用語를 制定하려면 語義分析이라는 매우 複雜한 問題의 고려가 先行되어야 한다. 語義分析의 첫 段階는 두 말 할것없이 辭典을 내는 일이다. 그리고 그 辭典의 定義들은 그 情報시스템에서 쓰기에 가장 適合한 形式으로 된 것이어야 할 것이다⁷⁾

辭典의 定義들은 그 시스템 안의 다른 낱말들과의 關係를 明白히 할 必要가 있으므로 綱目科屬식의 體系 속에 배열하여 어떤 用語가 더 一般의이며, 어떤 用語가 더 特殊한 것인지를 나타낼 수 있다. 로제이(Roget)의 디소오르스는 그 代表的인 것이다.

데이터·베이스·시스템에 있어서 用語問題는 그렇게 深刻한 것이 아니다. 오히려 二次資料檢索시스템에서 用語標準化의 必要가 더 절실하며 디소오르스가 各 別히 重要한 役割을 하게 된다.

예컨대 文獻內容의 記述은 여러 用語 中에서 선택한 主題語를 使用하여 行하는데 그런 主題語는 여러 가지 뜻으로 해석될 수 있으므로 그 시스템에서 通用되는 正確한 뜻이 어떤가에 固定되어 있어야 하는 것이다.

이런 固定을 위하여 主題標目典據와 디소오르스가 있는 것이며 이들은 결국 可用語들 中에서 뽑아 固定시킨 낱말을 보여주는데 지나지 않는 것이다. 이와 비슷한 意味로 볼 때 分類表라는 것도 用語들을 階層別로 배열하여 그들간의 關係를 表示한 一種의 사전이라 할 수 있을 것이다.

機械와의 對話問題

기계와의 대화에는 다음 네가지 사항이 考慮되어야 한다. (1) 사람과 機械가 다 같이 理解하고 使用할 수 있는 文字(記號), (2) 機械가 利用할 수 있는 範圍 안에 蓄積되어 있는 데이터·베이스, (3) 機械가 既存 데이터·베이스의 記號와 入力記號를 結合하고 對話할 수 있는 常用指示語, (4) 機械가 連續處理를 하게 하는 方法이다. 이러한 문제들은 自然語로 된 記錄物을 다루는 시스템에서는 特別 重要한 것이다. 그러나 방대한 用語量. 모호한 語義, 機械認識이 가능한 論理的 構文形式의 不足等으로 機械와의 對話問題는 아직도 엄청난 것으로 남아 있다.

索引, 抄錄, 記號化問題

情報시스템의 主要業務 가운데 하나는 여러곳에 흩어져 있는 文獻을 수집하여 그 內容을 索引, 抄錄, 또는 特定記號의 形式으로 記述한 후 시스템 속으로 入力하는 것이다. 이 때 각 자료 內容을 適切하게 表現하는 말들을 選定하고 그것들을 다시 記錄으로 綜合하는 方法의 開發이 必要한 것이다. 이는 실제로 어떤 자료내용을 分析하여 抽出된 各 主題에 해당되는 用語들을 用語辭典에서 選擇하여 賦與하는 행위이며, 바뀌 말하면 索引, 抄錄 등의 記錄을 만드는 과정상의 문제라 할 수 있다.

4.2 蓄積 및 檢索過程에서 일어나는 問題

세 問題領域으로 나누어 생각할 수 있다. 첫째 영역은 排列組織이며 이는 圖書館目錄카드의 順序를 정하는 排列法에 해당하는 것이다. 둘째는 利用者의 需要와 資料間의 合致度問題이고 마지막으로 圖書館參考奉仕에 해당하는 探索方法上的 문제인데 이는 書誌事項 등을 探索하고 評價하는 문제들과 關聯된 것이다.

5) Fodor, Jerry A., and Jerrold J. Katz, *The Structure of Language, Readings in the Philosophy of Language*, Englewood Cliffs, N.J., Prentice-Hall, 1964. Introduction.
6) Chomsky, Noam, "A Transformational Approach to Syntax," In Fodor and Katz, *op. cit.*
7) Salton, Gerard, *Automatic Information Organization and Retrieval*, New York, McGraw-Hill, 1968. Chapter 2 and 3.

排列組織

이것은 情報學者들이 오랫동안 매우 소홀히 다루어 온 分野로 이야기되고 있다. 다만, 방대한 蓄積情報을 效率적으로 利用해 보려는 몇몇 데이터 處理實務者들만이 排列에 대한 관심을 버리지 아니하고 分析的 研究를 계속해왔을 뿐이다⁹⁾. 이렇게 排列組織에 대한 研究가 소홀히된 理由는 아마도 이것이 다른 分野 特히 用語標準化에 얽힌 문제로 보아 온 경향때문인지 모른다. 하여간 排列組織의 基本問題들은 用語標準化와의 關聯을 인정하더라도 圖書館學 中心課題中 하나인것만은 틀림없는 일이다.

어떤 排列組織 방식이든지 먼저 排列順序가 정해져야 한다. 記錄들은 各各 어떤 位置를 잡아 거기 들어가서 蓄積되어야 하는 것이다. 排列組織에서 그 順序는 매우 重要하다. 特히 機械화된 경우 記錄의 位置는 후에 그것을 檢索하기 위하여 컴퓨터가 가야할 곳이므로 더욱 그러하다. 흔히 排列順序는 記錄樣式가운데 「記錄標識」라는 欄의 記入(예를 들면 文獻番號 같은것)에 따라 決定된다. 大部分의 圖書館에서 採擇하는 排列組織方式은 用語의 構造를 利用하는 것이다. 即 主題標目, 分類番號, 키어·워드 따위의 구조를 利用하는 것인데, 이 방식의 利點은 쉽고 單純하다는 것이다. 누구나 다 아는 방법이고 또 낯 익은 방법이다.

다른 排列組織法으로, 記錄內容의 類似性에 根據하여 組織하는 법이 있다. 이것은 비슷한 記錄들을 한데 모으는 방법인데 여기 제기되는 문제는 어떻게 유사성을 決定할 수 있는나는 것이다.

그러나 실상 主題標目이나 分類番號에 依한 組織도 잘 보면 共通 主題에서 오는 類似性에 依한 組織임을 알 수 있다.

需要와 資料 間的 合致度

情報시스템의 한 機能은 利用者의 要求에 合致되는 資料를 파일에서 가려내는것이다⁸⁾ 需要에 대한 어떤 資料의 合致與否는 利用者의 要請과 그 시스템의 選別過程을 通하여 決定된다. 選別過程의 最終選者는 물론 利用者이다. 利用者는 다 自己特有의 選別基準을 가지고 그것을 蓄積資料를 담은 그 시스템 特有의 形式에 適用함으로써 目的에 이르는 것이다. 그러나 利用者는 자신이 방대한 파일 全體에 대해 그렇게 할 수는 없으므로 그의 要請에 따라 豫備選別을 해서, 利用者 앞으로 올 때에는 合致도가 높은 몇개의 資料로 壓縮되어야 바람직하다 할 것이다.

要請은 百人百態로 提出되게 마련이다. 그 表現이나

主題分野, 要求하는 情報形式이 다 다를 것은 짐작하고도 남는다. 어떤 要請은 豫想需要에 對備하여 마련해 둔 資料에 依해 쉽게 답할 수 있을 것이고, 어떤 것은 情報選擇提供(SDI) 형식으로 계속 提供을 要하는 것일 수 있고, 또 어떤것은 면밀한 파일 調査를 要하는 것일 수 있을 것이다.

그러나 어느 경우에도 요청은 먼저 蓄積資料와 맞추어 볼 수 있는 形式으로 表現되어야 한다. 이는 요청 내용을 해석하여 必要한 記錄의 어느 欄에 記入되었을 用語(나 記號)로 表示함을 뜻하는 것이다. 完全한 要請을 提出하자면 利用者와 蓄積된 디소오터스사이의 對話가 必要하다. 兩者간 對話를 通해서만 利用者가 使用한 말들이 그대로 要請內容에 合致되는 파일 記錄을 指示할 수 있는 것이다.

要請이 일단 적절하게 形式化되면 그것에 合致되는 지 與否를 決定하기 위하여 파일 속의 모든 記錄은 일일이 調査를 받게 된다. 「合致度」라는 말은 多少 막연한 개념이므로 컴퓨터 프로그램에서는 예컨대 요청서 속의 낱말의 수와 蓄積된 기록의 낱말 수가 어느 정도로 비슷하느냐 하는 따위의 合致의 程度를 말한다. 그리하여 컴퓨터 檢索에서는 그런 種類의 숫적 유사성이 가장 높은 記錄을 「合致된다」고 본다.

큰 파일에서 모든 記錄을 일일이 要請과 比較한다는 것은 長時間을 要하여 能率的이라 할 수 없다. 이 문제의 해결자로 나타난 것이 合致도가 높은 記錄들의 位置를 알리기 위해 쓰이는 索引파일이다. 要請과 索引記錄들을 比較하는 일을 파일 探索이라 한다.

探索方法

파일 探索이란 具體적으로, 記錄들을 走査하며, 要請과 맞추어 보고, 合致되는 것들을 골라내는 過程을 말한다. 記述上的 잘못이나 不確實性때문에, 要請에 들어 맞는 記錄이 있어 과연 目的을 이루게 될것인지는 언제나 確實하지 않으므로 探索方法이라는 문제가 提起되게 마련이다. 어떤 파일에 대한 探索을 시작하기전에 探索方法 即 이 方向으로 할 것이냐 아니면 저 方向으로 바꿀것이냐를 먼저 決定해야 한다. 이 때 고려해야 할것은 所要時間, 所要費用, 豫想結果等이라 할 수 있다.

探索過程의 結果는 「合致되는 것」이라 하여 利用者에게 내주는 몇개의 檢索된 記錄일 것이다. 여기서 소

8) Meadow, Charles, *The Analysis of Information Systems*, New York, Wiley, 1967.

9) Cuadra, Carlos, *On the Utility of the Relevance Concept*, System Development Corporation, March 18, 1964.

探索 過程의 能率이 하나의 문제가 될 수 있다¹⁰⁾.

利用者가 어떤 情報시스템에 要請을 가져올 때에는 願하는 資料가 거기 있으리라는 期待를 가지며 要請의 응답으로 받은 자료는 요청내용에 합치된다는 保證을 바라는 것이다. 이 두가지 期待는 소위 「滿足率」「正確率」問題와 關聯되는 것으로 前者는 그 測定이 매우 어려운 것이나 後者는 비교적 측정이 容易한 것으로, 費用對效果調査는 그 한方法이라 할 수 있다.

4.3 情報提供過程에서 일어나는 問題

이것은 利用者의 요청에 응하여 데이터를 分析하고 情報를 提供하는 過程에서 일어나는 문제들이다. 먼저 一般的 需要를 充足시키기 위하여 마련하는 情報組織(記錄乃至出版)의 問題를 생각할 수 있으나 큰 문제가 아니며 다만 정해진 순서에 따라 情報를 배열하면 그만이다. 그 다음 문제는 개별 질문에 응답하는 일로서 이것은 너무 복잡하여 이의 기계화는 생각하기 어렵다고 보는 견해가 지배적이다. 두 문제의 學論이 다 機械화된 시스템에서만 의미가 있음은 물론이다.

情報組織

보통 파일 管理시스템에는 報告作成(report generation) 能力이 있다. 즉 出力(예컨대 索引 같은것)의 체계를 유지하고 배열순서를 맞추고 統計的 要約處理를 할 수 있는 것 등이다. 그러므로 檢索된 情報는 合致度가 높은 順으로, 時間順으로, 著者名의 알파벳順으로 등 어떤 順序로든지 배열할 수 있다. 代表的 체계는 KWIC 索引인데 이는 키·워드의 알파벳順으로 論文標題를 排列한 것이다.

質問應答

기계에 蓄積된 資料를 通해 論理的 推理를 하여 個別質問에 應答한다는 것은 매우 어려운 問題이다. 推論式, 微積分等을 應用한 實驗的 質問應答시스템이 試圖된 일이 있다¹¹⁾.

그러나 그 실험들이 너무 理論偏向의이고 실제문제에는 많은 制約點이 있어 실용가능성에 대해서는 많은 사람들이 悲觀的 見解를 表示하고 있다.

5. 結 言

오늘 날 圖書館利用者들의 要求는 變해가고 있다. 특히 科學技術分野讀者들의 要求가 그렇다. 그들은 特殊圖書館에서, 大學圖書館에서 「最近」 研究報告文獻을 찾아 血眼이 되어있다. 圖書資料보다 情報 그 자체를 찾는 경향이 늘어간다. 圖書館은 이런 새로운 要求의 內容과 형식에 부응하여 參考奉仕를 充實히 하는 한편 제어날 記事標題 서어비스, 書誌서어비스, SDI 서어비스 등 소위 커렌트·어웨어네스(current awareness) 방면 奉仕를 擴張하고 있다.

圖書館은 원하진 않건 間에 物體로서 圖書資料를 취급할 뿐 아니라 그 內容 即 情報를 處理(提供)하는 도큐멘테이션·서어비스에로 多角化하고 있다. 傳統的 圖書館은 변모하고 있는 것이다.

이런 狀況 속에서 圖書館人들은 強要되는 變化에 消極的으로 끌려가기만 할 것인가? 오히려 積極的이며 前進的인 姿勢로 바람직한 變化의 方向을 能動的으로 모색하고 도큐멘테이션과 情報學이 圖書館奉仕와 기술에 줄 수 있는 잠재적 利點들이 무엇인가를 캐어 알아야 할 것이다. 圖書館內에 情報奉仕를 위한 시스템도 우리 힘으로 設計해야 할 것이다. 이던것이 다 한꺼번에 쉽사리 成就될 일이 아니지만 筆者는 그 어렵고 먼 길의 스타아트 信號로 이 小論을 發表하는 것이다. 圖書館情報시스템의 세 主要過程 設計에 일어날 것이 예상되는 문제들을 한 二次資料 蓄積檢索 시스템 模型을 通해 要約해 보았다.

本論 속에서도 制限點으로 明白히 한 바와 같이 問題를 問題로만 要約했을 뿐 해결의 試圖는 하지 않았다.

이 分野에 關心있는 讀者들에게 研究의 方向 設定에 도움이 되었으면 했을 뿐이다. 그리고 앞으로 여기 提示된 問題들 하나 하나가 研究 發表되어 자주 本誌를 廣場으로 진지한 討論이 展開되었으면하고 바라는 마음 간절하다.

10) Lancaster, F. W., *Information Retrieval Systems Characteristics, Testing, and Evaluation*, New York, Wiley, 1968.

11) Hayes, Robert M., "Review of Asa Kasher, *Data Retrieval by Computer*", *American Documentation*, vol. 18, no. 3 (July 1967), 187-89.