UDC 681.142.8—532.8

Automatic Documentation, Part 1

MIRACODE와 FLIP

Unesco 科學文獻專門家 **Dr. Walter T. Lorch**
KORSTIC 事務總長 **吳 益 相**

SUMMARY : This first article in a series of descriptions of devices and methods in the field of automated information processing depicts two machines representing a certain type of documentation systems which use micro-images and coding of documents in an unalterable sequence, namely on 16mm filmrolls.

科學情報의 蓄積과 檢索方法은 Documentation과 Aspect의 두 方法으로 나눌 수 있다.

「Documentation方法의 特徵은 文獻 또는 그 一部分을 寫眞으로 蓄積하였다가 code로서 檢索하는 것이다. 이 code는 該當文獻의 寫眞(때로는, 抄錄이나 文獻의 一部分 또는 設計圖만의 寫眞) 귀퉁이의 感光層에 透明 不透明 點의 配列로 構成되어 있다. Documentation方法에는 數千의 文獻과 對應되는 code를, microfilm두루마리에 一定不變順序로 固定蓄積시키는 FLIP이나 MIRACODE方法과 個個의 microfilm에 可變順序로 蓄積하여 두는 FILMOREX나 MINICARD方法의 두 型이 있다. code로서 記述特徵 지워진 이러한 어떤 文獻을 檢索할 때, 첫 方法으로는 하나 또는 몇개의 microfilm 두루마리 全體를 選擇機에 걸어 順序대로 죽 훑어 보아야 되는 反面 둘째 方法에 있어서는 처음에 蓄積할 때 個個의 microfilm을 어느程度로 細分하여 分類하여 두었느냐에 따라 달라지기는 하지만 훑어 보아야 할 microfilm의 數는 먼저 方法보다는 훨씬 작다. 따라서 檢索速度와 microfilm의 磨損에 對한 保存度의 兩面에 있어서 둘째 方法이 처음것 보다 優秀하다고 생각된다.

그러나 本稿에서는 microfilm두루마리(reeled film rolls)를 使用하는 FLIP과 MIRACODE方法에 依한 科學情報의 蓄積과 檢索의 問題를 다루어 보기로 한다. 이 두 方法 모두 美國에서 考案된 것이다.

FLIP 方法

"Film Library Instantaneous Presentation"의 略稱에서 나온 이 FLIP方法은 1958年 Benson-Lehner株式會社(11930 West Olympic Boulevard, Los Angeles 64, California)에서 開發된 것이다.

film의 各 토막은 다음의 두 部分으로 構成되어 있다. 即 文獻의 寫像과 이 film寫像의 右側 맑은 背面에 黑色의 토막 줄로 構成된 總 32個의 二進法 code單位이다. 兩側으로 穿孔된 폭 16mm의 film에 이 두 部分이 다 들어 있다. 各 microfilm두루마리의 film 길이는 1,200회—트이며 數百부터 最大 72,000 film토막을 包含할 수 있다. 두루마리의 直徑은 12인치이다. 文獻의 蓄積量은 좀더 큰 두루마리에 film을 많이 감으면 쉽게 增加할 수 있으나 必然的으로 부피가 커져서 檢索時間이 길어진다는 短點을 가져온다. 그렇다고 두루마리의 film길이가 너무 짧아도 안된다. 그것은 選擇機에 걸어 감고 다시 감어 놓는데 걸리는 時間이 增加되기 때문이다.

緻密하게 構成된 microfilm의 讀字 및 檢索機의 무게는 約 900파운드이다. 鍵盤에 檢索할 情報를 指示하고 始動자루를 누르면 機械는 두루마리를 回轉시켜 要求하는 情報를 담은 film토막이 發見될 때까지 自動的으로 film을 走査한다. 檢索速度는 每秒 길이 60인치의 film을 또는 300個의 film토막을 走査하는 程度이다. 이速度는 좀 느린 便이라고 생각된다. 走査時 이 選擇機의 感光要素는 film토막의 文獻寫像 귀퉁이에 記錄된 黑白의 二進法數字 code를 읽게 된다. 檢索하려던 film토막이 發見되자마자 그것은 스크린에 投影되고 同時에 film토막 選擇의 立證으로서 指示器에 code數字의 一致가 燈示된다.

各 film토막에는 二進法數字가 8列에 4個씩 配列되어 있으며 이들 總 32個數字의 適切한 組合은 檢索에 必要한 情報를 記述하지 않으면 안된다. 그러나 8列에서 制約되게 된 8個의 記述子만으로는 文獻의 內容이나 型態를 充分히 記述하지 못할것으로 보이며 約 倍의 記述子가 必要할 것으로 생각된다. 따라서 이 機械의 檢索能力은 相當히 限定되어 있다고 하겠다. 이러한 難點은 한 film토막에다 文獻과 code를 記憶시키지

말고 두 film토막을 使用하여 한쪽은 文獻을, 다른 쪽은 索引 code만을 利用한다면 쉽게 克服되나 한 microfilm두루마리의 情報蓄積容量을 減少시키며 따라서 檢索速度도 떨어지게 되고 또한 機械도 相當한 改造를 입지 않으면 안된다(한 例로서 機械의 感光細胞素數도 增加된다). 本機械는 約 한 時間동안에 約 50萬個의 文獻을 走査하여 該當 文獻들을 檢出해 낼 수 있다. 그러나 이 檢出된 文獻들은 比較的 적은 數의 記述子로 檢索된것으로서 좀 넓은 分類項目에 屬하는 것들이기 때문에 사람이 最終的으로 檢査하여 質問에 가장 適合한 것을 고를 必要가 있다. 이 機械의 製造會社 說明書에는 한 時間동안에 1,080,000個의 文獻을 檢索할 수 있다고 되어 있는데 이것은 여러 microfilm두루마리 中에서 가장 適合한 두루마리를 찾는 時間과 그것을 機械에 걸고 그 後에 다시 감는 時間 等은 考慮外에 둔것이라 하겠다.

該當 film토막을 發見後 機械는 다섯토막 좀 더 돌아갔다 다시 되돌아와 該當 film토막을 光學機械 앞에 멈추게 한다. 이것은 스크린에 그 film토막의 擴大像을 投影한다. 普通 室內의 밝기에서 解像力의 損失은 別로 없다(1mm 間隔內의 25線에 對한것보다 낫다). 이것은 다만 읽을 수만 있을 뿐이지 投影된 像을 複寫할 수 있는 準備는 되어 있지 않다.

MIRACODE 方法

Microfilm Information Retrieval Access Code의 略稱인 MIRACODE方法은 類似한 原理 밑에서 이루어 진 것이나 最近의 이 分野의 發展을 나타내고 있다. 普及되고 있는 이 機械의 製造處는 Recordak株式會社(415 Madison Avenue, New York 17, N.Y.)이다.

蓄積: 첫 段階는 microfilm으로 될 文獻을 調査分析하여 어느 記述子가 文獻과 그 內容을 가장 잘 鑑定해 낼수 있느냐를 決定하는 것이다. 이것을 위해서 記述子와 그에 對應되는 세자리數로 構成된 code를 收錄한 一種의 指針書가 利用된다. 이것은 各各의 主題部門에 있어서 999個 까지의 記述子를 利用할 수 있다는 것을 意味한다. 普通 한 文獻內容을 code化하는데에는 12個 乃至 15個의 記述子가 必要하다. 그러나 檢索指示用으로는 훨씬 적은 數의 記述子가 使用된다. 即 檢索指示用으로는 普通 5個 乃至 7個의 記述子이면 充分하다. 境遇에 따라서는 文獻內容을 表示하는 記述子를 使用하지 않고 記錄番號나 計劃番號같은 文獻의 整理番號를 使用한다.

이와같이 調査된 記述子의 code數를 (세자리의 數로 構成되어 있는) 普通 세줄로 나누어 所謂 code sheet에 記錄한다. code sheet는 $8\frac{1}{2}''\times11''$의 標準形 便紙紙의 크기다.

이 code sheet와 該當文獻은 16mm microfilm에 轉寫되게 된다. Miracode Microfilmer는 이 目的을 위해서 特히 考案된 附隨裝置로서 高解像力, 正確한 照明, 操動의 便利性 等의 特性을 지니고 있다. 文獻과 該當 code는 $\frac{1}{24}$의 縮小率로 길이 100휘—트의 film두루마리의 16mm토막에 microfilm化한다.

그러나 文獻은 그대로 microfilm化 되지만 code sheet에 있는 code數字는 그대로 寫眞化되는 것은 아니다. 이들 code數字는 二進法 數字 code로 變更되어 透明不透明의 短形무늬로서 microfilm에 나타나게 된다. 이 矩形무늬의 配列이 個個의 文獻을 識別하는 索引 code인 것이다.

이와같이 索引 code를 film에 옮기기 위해서, 色으로 쉽게 區別할 수 있게 되어 있는 數字選擇 스라이드 스윗치가 9個 있다. code sheet에 記錄되어 있는 數字대로 스윗치만 누르면 二進法 code化된 무늬가 film에 옮겨 진다. 正確히 code化 되었나를 檢査할 수도 있다.

MIRACODE microfilm두루마리의 情報蓄積容量은 FLIP의 約 $\frac{1}{3}$로서 한 文獻의 平均 page를 6 page라고 본다면 이 100휘—트의 film두루마리에 2,000 page의 文獻이 索引 code와 같이 蓄積된다.

이와같이 處理를 끝낸 100휘—트의 film두루마리는 各各 4″×4″×1″의 film들(箱子)에 保管된다. 이 film들은 主題分野別로 區分되어 쉽게 손이 닿을 수 있는 선반에 整理된다. 이 선반은 490個의 film들을 쉽게 손이 닿을 수 있는 位置에 保管할 수 있다. 또한 그 配列方法도 여러가지가 있다.

檢索: 特定의 情報에 對한 要請이 들어오면 蓄積때 使用하던 그 code索引 指針書를 利用하여 먼저 그 質問內容을 二進法 code로 옮겨야 한다. 主題記述子의 配列順序는 檢索過程에 있어서는 相關없다.

檢索者가 質問內容을 너무 細密히 分析하여 너무 많은 記述子를 가지고 檢索하려고 할 때 該當文獻을 찾지 못하는 境遇가 종종 일어난다. 이때에는 덜 重要한 記述子로 부터 차츰 줄여 가면서 結果를 얻을 때까지 檢索은 反復한다. 檢索索引書는 어느 들에 問題의 記述子를 包含하고 있는 microfilm이 들어 있는가를 指摘한다. 그러면 그 microfilm들을 끄내서 檢索機에 걸면 自動的으로 走査하면서 高速檢索을 끝내고 다시 原狀대로 감긴다.

Recordak Lodestar Reader-Printer, Model PEK는 microfilm의 二進法 code를 電子的으로 高速 檢索하며 檢索해 낸 文獻을 印刷까지 한다. 陰畵 또는 陽畵 microfilm을 檢索하는데 必要한 操縱者의 모든 作業까지도

해낼 수 있는 鍵盤制御器에 이 機械는 連結되어 있다. 平均檢索時間은 넌지 8秒에 지나지 않으며 이것은 操縱者가 檢索指令단추를 누른 瞬間부터 該當 film을 찾어 스크린에 그 寫像을 보이게 하거나 또는 印刷段階에 들어갈 수 있게 할 때까지 걸리는 時間이다.

code化된 microfilm은 每秒 10회—트의 率로 走査된다. 即 FLIP方法 보다 約 두倍의 速力이다. 質問內容에 適合한 文獻의 첫장이 發見되면 film走査는 自動的으로 멎게 되고 그 像이 스크린에 投影된다. 이 以外에도 같은 質問條件을 滿足하는 다른 文獻을 찾으려면 다시 檢索指令단추를 누르므로서 microfilm들의 나머지 部分에 對한 檢索이 繼續된다. 이와 같이 900,000 page 以上의 文獻中에서 무엇이든 15秒內에 檢索해 낼 수 있다.

情報를 要求한 者는 必要한 情報를 스크린에 비쳐진 影像이나 또는 종이에 複寫된 形態로 얻을 수 있다. 또 手動的 複寫方式以外에 自動的인 네가지 複寫方式中 어느 하나를 擇할 수도 있는 것이다.

結　論

FLIP과 MIRACODE의 兩方式을 比較하여 볼때 情報管理의 機械化에 있어서 단지 數年동안에 이루어진 急速한 發展에 驚嘆하지 않을 수 없다. 이미 莫大한 量의 文獻이 잘 分類되어 있고 그 增加速度가 느릴 境遇에는 MIRADODE方法은 大端히 效率的이다. 文獻의 microfilm像과 該當 code가 서로 物理的으로 連結되어 100회—트의 film두루마리를 構成하고 있어서 走査速力은 比較的 높다. 勿論 좀 더 긴 film두루마리를 使用할 수도 있으나 이때에는 좁은 主題分野別로 整理되어 있는 microfilm들을 使用할 수 있는 利點을 잃게 된다. 固定不變의 film順序를 지니고 있는 이 두 方式의 欠點은 蓄積의 처음 段階에 使用하였던 分類方法을 끝내 使用하지 않으면 안되며 나중에 바꿀 수 없다는 데에 있다. 情報蓄積에 있어서 이러한 順序의 凍結은 檢索時 不得已 一部情報의 漏落이 不可避하다. 不然이려면 선반의 모든 microfilm들을 選擇機에 걸어 檢索하는 努力을 들여야 한다.　　(吉炳敏 譯)

(p.1에서 繼續)

調 및 事務當局의 쉴 사이 없는 努力으로 많은 情報를 蒐集하여 供給하고 또 能率的으로 그것을 管理하여 많은 科學者들에게 利用될 수 있게 한 까닭에 好評을 받게 되었으니 또한 기쁜 일이라 아니할 수 없다.

過般 서울大學校內에서 開催된 第6回 原子力學院會議 第四分科 심포지움 "科學技術行政體制問題에 對하여"라는 題目을 다루는 가운데서 科學技術의 綜合的인 研究事業을 爲한 科學技術開發公社를 新設하려는 意見이 있었다 하는데 이 亦是 우리나라 科學研究를 爲하여 當然히 있어야 할 機關이라고 본다. 이것이 國營研究機關이 되든 또는 民間研究機關이 되든 KORSTIC와 더불어 우리 韓國科學研究의 스폰서—로서 우리나라 科學의 振興과 產業의 原動力이 되기를 祈願하며 아울러 科學者 여러분의 愛用과 愛護를 얻어 앞으로 KORSTIC이 이나라의 文化 및 生產向上에 더 큰 貢獻이 있게 되기를 빌어마지 않는다.

(p.5에서 繼續)

를 收錄하게 되면 2000年에 發刊될 第72卷은 抄錄記載面이 120,000頁, 主題索引面이 21,000頁이나 될 것이며, 그 두께는 30회트나 되어 書架의 10미터를 차지하게 된 것이다.

오늘날 人間知識의 寶庫에는 每年 約 6百萬의 科學論文이 入庫되고 있다. 그리고 이 많은 論文은 年間 10~20%式 增加한다. 따라서 科學者 個個人이 自己專門分野의 情報를 自力으로 收集하던 在來方法은 낡은 手法이라기 보다는 不充分한 方法이 되어버렸다.

現代科學者들이 最新知識에 밝고, 重複된 研究를 避하자면 보다 迅速正確한 方法을 講究하여 最新情報를 入手하여야 할 것이다. 이 問題에 對한 解決策으로 登場한 것이 Documentation이고 Documentation에 있어서 새롭고 革新的인 發展을 보았다면 그것은 바로 科學情報處理의 自動化이다.　　(申鉉斐 譯)