

K.I.E.T. 經濟統計

데이터 베이스에 관한 小考

韓 泰 仁
(電算室)

..... 〈차 례〉

- I. 目 的
- II. 데이터 베이스의 構成
 - 1. 構 成
 - 2. 資料의 種類
 - 3. 資料의 運用
 - 4. 데이터 베이스의 利用方法
 - 5. 國內 綜合經濟情報 體制와
의 連結
- III. 結 論

I. 目 的

社會 現象에 관한 計量的 研究方法이 急増함에 따라 수 많은 人間과 社會活動에 대한 統計資料가 해당分野의 研究에 利用되었으며, 더우기 다른 社會部門과 달리 經濟的 現象에 관한 計量分析은 매우 急速하고 多樣하게 發達되었고 여기에는 産業, 人口, 貿易 및 經濟一般에 관한 共通的이고 普遍的인 基礎資料가 많이 使用되었다.

물론 이러한 基礎資料의 蒐集은 우리나라를 例로 볼 때, 經濟企劃院調查統計局이나 韓國銀行 같은 一國의 特定機關 내지는 이 分野에 特別한 研究를 담당하는 여러 研究機關으로부터 이루어지며, 蒐集된 資料는 안으로는 國內의 各 企業 및 機關의 研究 또는 單純한 情報로서 提供되어지며, 밖으로는 國家相互간 情報

交換 또는 情報販賣의 方法으로 他國의 基礎統計資料가 活用되어져 왔다.

그러나 情報의 交換 또는 購入이 可能하다고 하더라도 管理 및 普及의 問題는 情報의 活用面에서 볼 때 훨씬 重要的 部分으로 考慮되고 있다. 이러한 問題를 살펴보면 우선 資料의 手入先을 單一化하여야 할 것이고, 管理와 저장의 反復을 줄여야 할 것이며, 나아가서는 이렇게 하여 저장된 많은 資料에 쉽게 接近하여 빠르고 容易하게 願하는 質과 量의 資料를 얻어 낼 수 있는 資料의 檢索方法體系(Retrieval Method System) 역시 構想되어야 했다.

이러한 目的아래 이루어진 經濟데이터베이스는 世界各國에 서로 다른 이름으로 많이 散在해 있다. 프랑스의 GSI-ECO, G.CAM Questel 이나 캐나다의 I.P. Sharp, 獨逸의 INKA, BRS의 HARF, FNTL, FSIS¹⁾ 등 無數히 많은 經濟데이터베이스가 있다. K.I.E.T.의 ESDB²⁾ 역시 1984년 6월 같은 目的으로 OECD(經濟協力開發機構)³⁾ 와 IMF(國際通貨基金)⁴⁾ 그리고 國內의 一般經濟統計를 電算機에 收錄하여 온라인(on-line), 對話式으로 檢索할 수 있는 經濟統計 데이터베이스로 製作되었던 檢索시스템이며 1985년 1월에는 UN(國際聯合)⁵⁾의 産業別 資料와 國民計定資料를 追加 補完하였다.

II. 데이터 베이스의 構成

1. 構 成

ESDB의 시스템을 構成하고 있는 主要諸元을 보면 첫째, 機械의 種類로는 IBM 4341의 本體와 온라인용으로 使用하고 있는 端末機가 있다. 이 端末機는 한글, 한자 및 英字 兼용 端末機로서의 NEC 6300/50 N⁶⁾을, 英字만을 위한 것으로는 IBM 3278 CRT⁷⁾를 주로 使用하고 있다. 둘째, 온라인 對話式 檢

1) 「主要情報源 紹介」, 産業研究院, 1984.11.

2) Economic Statistics Data Base.

3) Organization for Economic Cooperation and Development.

4) International Monetary Funds.

5) United Nations

6) National Electric Company의 Display terminal

7) Cathode Ray Tube.

索을 위한 소프트웨어(software)에는 DOS/VSE⁸⁾ 시스템과 CICS⁹⁾ 프로그램, PL/I¹⁰⁾ 컴퓨터言語(computer language)를 사용하였으며 세계, 자료로는 K.I.E.T. 데이터뱅크(data bank) 중에서 産業經濟統計部分의 資料를 選擇하여 構成하였다. K.I.E.T. 의 데이터베이스에는 電子, 化學, 의학 등의 技術정보와 特許情報 및 海外 데이터베이스가 있는데 그 중에서 經濟資料에 關連된 一般經濟와 産業情報의 資料를 選擇하고 貿易에 關連된 國別 輸出入統計를 收錄하지 못한 것은 電算化하기에는 資料가 너무 방대하기 때문에 데이터베이스의 構成資料에서 제외되었다.

2. 資料의 種類

(1) OECD 資料

國際經濟協力開發機構의 資料는 會員國 24 國(參觀國: 터어키, 유고包含)에 對한 國民所得, 産業別生産, 運輸, 通信, 國債, 換어음, 建築, 企業經營, 雇傭, 賃金, 物價, 財政, 國際收支, 貿易 등에 關한 統計資料이다.

(2) IMF 資料

國際通貨基金의 資料는 會員國 175 個國에 對한 經濟統計資料로서 크게 IFS¹¹⁾ (國際財政統計), BOP¹²⁾ (國際收支統計)와 GFS¹³⁾ (國家財政統計)로 나눌 수 있으며, IFS에서 換率, 通貨金融機關, 通貨概觀表, 流動性概觀表, 預金銀行, 金利, 物價, 生産, 國際去來, 財政 및 國民計定統計를 收錄하였고, BOP에는 經常收支計定, 直接投資와 여러 部門의 長短期 資本收支, 財貨, 用役 및 預金支給準備 등의 資料가 있으며 끝으로 GFS統計에는 豫算計定, 總括計定, 特別豫算計定과 社會保障基金 등 中央政府나 州政府 혹은 地方政府에 對한 一般計定을

8) Disk Operating System / Virtual Storage Extended.

9) Customer Information Control System.

10) Programming Language I.

11) International Financial Statistics.

12) Balance of Payments.

13) Government Financial Statistics.

다루었다.

(3) UN資料

國際聯合의 資料는 GIS¹⁴⁾와 CPS¹⁵⁾ 그리고 NAS¹⁶⁾ 統計로 나뉘는데 GIS와 CPS는 主要産業別 各種 産業經濟統計 및 商品別 産業統計를, NAS에서는 國內生産에 관한 最終支出, 一般政府, 法人 및 準法人企業, 家計와 民間 등과 같은 機關別計定, 業種別 生産 및 計劃經濟에 관한 資料를 收錄하고 있다.

(4) 國內 經濟統計

國內의 資料에는 産業別 生産에 관한 많은 資料 中에서 生産指數, 生産出荷指數, 生産者製品在庫指數에 대한 原指數 및 季節調整指數가 있으며, 生産者出荷量과 製品在庫量, 生産能力指數와 製造業 稼動率指數 등 産業別 主要統計와 國民總生産과 産業別 國內總生産 및 國民總生産에 대한 支出 등 國民計定에 관한 一般的인 經濟統計와 全都市에 대한 品目別 消費者物價指數, 都賣物價指數, 都賣價格과 輸出入物價指數 등 物價에 관한 全般的인 資料를 收錄하고 있다.

3. 資料의 運用

國內外로부터 入手된 資料는 앞에서 살펴 본 바와 같이 經濟統計分野라는 것 外에는 서로 共通的 事項이 없이 그 形態나 內容이 相異하다. 물론 各 資料에서 취급되고 있는 對象이 一定하지 않으며, 같은 內容의 資料가 있다고 하더라도 그 分類의 目的이나 方法 등의 異見에 따라 다른 모양으로 다루어지기도 한다. 그러나 그보다 더 根本的인 問題는 基礎資料의 製作機關이 같지 않음으로 해서 發生되는 製作過程上의 諸般問題이다.

또한 經濟統計資料의 內容上으로 볼 때 資料生成의 發生頻度와 發生時點 역시 또다른 問題點을 지니고 있다. 우선 어느 電算機든 記憶用量을 생각하지 않을 수

14) General Industrial Statistics.

15) Commodity Production Statistics.

16) National Account Statistics.

없기 때문에 資料의 選擇을 신중히 해야 한다. 예를 들어 매일매일, 그것도 各業體別로 發生하는 證券時勢 같은 經濟資料는 매우 收錄하기 어려울 것이다. 뿐만 아니라 한 資料를 購入하여 使用하는 過程에 비해 資料의 生成時間이 많이 소요되고, 항상 最近의 資料만을 必要로 하는 경우 이 資料 역시 選擇의 여지에서 考慮되어야 할 것이다.

이와 같이 資料의 運用面에서 보면 어려운 점이 많이 있는데 이를 要約해 보면 다음과 같다.

첫째, 購入資料의 供給源이 一定하지 않기 때문에 資料의 形態가 매우 多樣하다. 따라서 이들을 손쉽게 다룰 수 있도록 資料를 處理하여야 한다.

둘째, 資料의 內容이나 形態上 그 種類가 매우 많으므로 購入先의 選擇도 必要하고 더우기 電算機의 構造를 감안해야 하므로 그 많은 資料 中에서 가장 普遍的이고 多樣하게 使用되는 一般的인 經濟資料를 選擇하여야 한다.

셋째, 資料使用의 特性 내지는 그 內容上 時間의 민첩성을 要하는 資料가 있다. 即 物價나 物價指數에 관한 資料처럼 資料의 入手도 빨라야 하며 入手되는 대로 使用可能하도록 하여야 하는 어려운 資料도 있다. 이 이외에도 많은 條件들이 있겠지만 ESDB에서는 주로 이러한 問題點을 考慮하여 데이터베이스의 資料를 運用하는 方法으로 購入過程과 期間 및 資料의 소장량을 決定하였다.

各 資料에 대한 購入過程 및 소장량을 비교해 보면 첫째, UN資料의 경우 國民計定을 다루는 NAS와 各産業別, 品目別指數 및 一般經濟統計에 관한 CPS와 GIS 統計資料는 거의 全世界가 對象國이므로 1년에 1회씩 資料가 生成되어 購入上の 問題는 없다. 그러나 品目は 事實上 品目群에 해당하는 中分類 以上の 코드로 축약해서 全品目を 國家別로 所藏해야 했다. 따라서 약 160,380 系列의 資料를 1년에 1회씩 修正補完하고 있다.

이와는 달리 둘째, OECD와 IMF資料의 경우에서 보면 各機關은 每月 資料를 補完하고 있다. 물론 資料의 發生이 年別, 分期別 資料인 것은 해당 時點에 따라 修正補完이 된다. 現在까지의 使用頻度數로 보아 IMF의 175國에 대한 147,900 系列의 資料가 가장 많이 쓰이며, 換率이나 物價 및 收支部分은 時間性を 要하는 경우가 많다. 그리고 OECD資料는 各國의 政府에 관한 一般統計로서 약 5,530 系列의 資料가 收錄되어 있다.

마지막으로 國民計定, 物價, 産業別 品目別生産에 관련된 資料를 處理하는 國

內經濟統計는 약 5,100 系列이 있으며, 이 資料가 ESDB의 內容中 가장 빠른 資料處理를 要하며, 實際로 資料의 發生과 購入이 매우 빠른 편이다. 이 資料역시 年別, 分期別 資料는 해당時點에 補完이 되고, 産業에 관련된 資料는 細細分類까지 되어 있으므로 구체적인 品目까지 그 基礎資料를 파악할 수 있다. 그러나 月別資料는 80년 以後의 것만을, 分期別資料는 70년 以後의 것만을, 年別資料는 53년 以後의 資料를 收錄해 電算機와 資料의 量과의 構造問題를 適切히 處理할 수 있도록 計劃設計되었다.

4. 데이터베이스의 利用方法

ESDB는 基礎資料의 온라인檢索을 위해 製作된 經濟統計 데이터베이스이다. 따라서 端末機를 통하여 所藏되어 있는 各種資料에 接近하여야 하므로 약간의 시스템利用方法을 必要로 하고 있다.

다른 모든 電算處理의 作業에서와 같이 이것을 3段階로 나누면 端末機와 ESDB시스템과를 連結하는 過程과 資料의 檢索過程 그리고 端末機와 시스템과를 分離시키는 過程으로 나눌 수 있다.

우선 첫번째 段階를 보면, 端末機의 種類에 따라 本體와 端末機를 連結하는 方法은 多少 差異가 있으나 이는 字板의 글자배열의 問題일 뿐이다. 그리고 이 시스템과의 連結은 CICS 프로그램에 의한 "ESDB" 라고 하는 transaction ID를 打字함으로써 이룰 수 있다. 두번째 段階를 살펴보면 우선 內容이나 形態上으로 資料의 種類가 매우 多樣하므로 各 資料가 만들어지는 資料群을 選擇해야 하고, 國家를 選擇해야 하며, 時系列의 區分과 願하는 資料의 正確한 變數名을 찾아 資料에 接近해야 한다. 이때 이러한 내용들을 端末機의 畫面에 打字하여 本體로 보내며 그 結果는 端末機의 畫面으로 出力된다. 따라서 이를 제공받으려면 하드카피의 方法을 利用하여야 한다. 이 외에 利用者의 便宜를 위하여, 구하여진 各種 資料의 소수점을 변경하는 것과 題目의 表現 등을 自由롭게 할 수 있는 프로그램이 內在되고 있다. 끝으로 端末機와의 分離인데 이는 各 資料群을 選擇하는 過程에서 어느 資料도 選擇하지 않는 경우 經濟데이터베이스로부터 分離되도록 設計되었다. 이때는 물론 端末機와 本體는 連結이 되어 있으므로 다른

作業을 할 수 있게 된다.

5. 國內 綜合經濟情報 體制와의 連結

國內의 各機關이 個別的으로 여러 種類의 資料를 소장하기 때문에 發生되는 情報管理上의 問題點으로 첫째는 같은 資料의 反復的 入手이고, 둘째는 資料의 加工에 소요되는 人力 및 豫算의 낭비이며, 셋째는 이들의 管理, 運營과 普及의 重複일 것이다. 따라서 이러한 폐단을 없애고, 效率的으로 統計資料를 管理, 利用하기 위하여 經濟企劃院 調查統計局, 韓國銀行, 韓國開發研究院, 貿易協會와 産業研究院이 自體所藏資料를 中心으로 해당機關에 맞는 基礎資料를 入手하여 手入先을 單一化하고, 研究方向 및 目的에 알맞는 소프트웨어를 開發하도록 하여 基礎資料는 물론, 適切하게 加工된 內容의 資料도 相互 交換하여 使用이 可能하도록 하는 綜合經濟情報體制를 構築하기에 이르렀다. 現在는 電算機의 하드웨어 構造上 調查統計局, 韓國開發研究院 그리고 産業研究院만이 이 體制를 이루고 있기는 하지만 나머지 機關역시 곧 關與할 展望이다. 어쨌든 이러한 體制는 個別的으로 固有한 資料를 分擔하여 入手, 所藏, 管理하면서 情報網(network)을 통하여 어느 資料든 쉽게 使用할 수 있도록 한 體制라는 點에서 業務開發의 效率化나, 資料의 利用極大化 및 人力과 財源의 蓄積面으로 상당한 役割을 하고 있다고 判斷된다.

Ⅲ. 結 論

위에서 살펴본 바와 같이 資料를 處理하여 데이터베이스를 構成하는 데는 많은 어려움이 따른다. 그러나 이보다 더 큰 어려움은 資料를 어떻게 處理하여 利用者에게 提供함으로써 이 資料의 利用을 便하게 하는가 하는 方法이다. 즉 利用者が 經濟데이터베이스를 接할 때에 가장 必要로 하는 것이 어떤 資料인가를 念頭に 두지 않을 수 없다. 이러한 問題는 다음 세가지 측면으로 생각해 볼 수가 있다.

첫째, 固有의 基礎資料를 利用者が 願하는 형태대로 檢索할 수 있어야 한다.