

情報處理의 標準化*

柳 京 熙**

1. 머리말

電子計算機가 이른바 情報나 데이터를 迅速하게 處理해 낼 수 있는 能力을 가지고 있음은 周知의 事實이다. 그러나 電子計算機는 約束된 電氣的 信號에 따라서 實行하도록 되어 있으므로 電子計算機를 使用하는 사람끼리 “어떻게 約束” 할 것인지를 미리 統一해 두어야 할 必要가 있다. 가령, 어떤 곳에서는 男性을 1로 하고 女性을 2로 約束해서 情報蓄積을 하고 있고, 다른 곳에서는 男性을 0으로 女性을 1로 約束했다고 하면 이 두 시스템 間에는 性別에 관한 情報에 관해서는 互換性이 없다. 또한 A機關의 物品管理를 위한 物品코드(Code)는 BTN 分類基準에 따르고, B機關에서 SITC基準, C機關에서는 US-FSC基準에 따라서 各各 約束이 되어 있다면 各各의 蓄積情報는 互換性이 없을 뿐 아니라 더욱 深刻한 問題가 생기게 된다. 즉 同一情報가 2重3重의 다른 方式으로 處理되므로 莫大한 經費가 낭비되는 結果를 招來한다는 것이다.

國內에서는 이미 이 分野의 教育機關이 많이 있으나 教授하는 用語가 모두 달라 專門家 養成에 있어서 滿足할만한 教授方法이 없는 實情이다. 그러므로 用語의 統一作業도 서둘러야 할 課題中的의 하나이다.

現在 우리나라에 導入 稼動中인 電子計算機는 86基에 達하며 年中 100基를 넘어설 것으로 推定된다. 數字로 보아서는 1萬基가 넘는 先進國과 比較도 할 수 없는 實情이지만 效率的이며 經濟的인 處理를 위한 바탕을 講究해 두지 않으면 안될 段階에 이르렀다고 본다. 이러한 目的을

두고 情報處理의 標準化事業이 段階的으로 推進되어야 할 것이다.

2. 標準化의 対象 및 範圍

情報處理의 標準化 對象을 綜合해 보면 다음과 같다.

表 1. 情報處理關係 標準化의 対象

區分	對 象
소프트웨어 관계	用語, 表現, 코우딩, 코우드, 파일 및 레코오드, 프로그래밍言語(ASSEMBLER, COMPILER, FORTRAN, ALGOL, COBOL, PL/1, RPG) 서어비스 프로그램, 제어프로그램(오퍼레이팅시스템, 制御프로그램, 멀티프로그램) ON-LINE 시스템(制御, 入出力, TIME SHARING 시스템) 情報시스템(情報檢索 파일) 應用프로그램(數理, 行列, 統計, LP, 시뮬레이션, 管理시스템, 自動制御, 數值制御, 圖形處理, ID, IE, RP, 시스템設計) 도큐멘테이션 소프트웨어 開發(프로그래밍 및 프로그래밍 프로젝트)
하드웨어 관계	시스템구성 (ARCHITECTURE, CPU) 論理裝置(回路 및 回路素子) 入出力裝置(磁氣테이프, 종이테이프, 디스크카드, 드럼, 타이프라이터, 穿孔裝置, 프린터, 플로터, 디스플레이) 其他 端末機器
통신시스템 관계	通信網(電話網, 回線網, 信號方式, 接續, 傳送, 家內裝置) 데이터 傳送(通信網, 回線分岐 및 交換方式, 變復調裝置 및 INTERFACE, TIMING 制御, ERROR 制御, 直列·並列 傳送) 端末機器 시스템構成 (ON-LINE, OFF-LINE 시스템, REAL TIME, SEMI-REAL TIME 시스템, 信賴性, 通信制御, MODEM)

*이 事業은 科學技術處의 用役事業에 의하여 遂行된 것임.

**KORSTIC 電子計算室長

以上の對象에서 가장 時急하다고 認定되는 것은 소프트웨어 分野로서 특히 用語 및 데이터 코우드 (Representation of Data Element)의 統一化로 들 수 있다.

本論에서는 데이터 코우드의 統一化를 위한 基礎資料의 提供에 重點을 두어서 言及하기로 한다.

3. 데이터 코우드

데이터 코우드란 情報의 單位가 되는 데이터 項目을 使用目的에 따라 識別, 分類, 排列하기 위하여 數字, 文字 또는 그 외의 記號로 表現한 것으로서 基本的으로는 識別하는 機能, 分類의 機能 및 排列의 機能을 갖는 3種類의 코우드가 있는데 코우드의 設計 以前에 이것을 충분히 理解해 둘 必要가 있다.

● 識別을 위한 코우드

한 개의 일이나 概念에 대하여 1대1로 對應하여 붙여지는 코우드로 어떤 것을 區別하기 위한 目的으로 쓰인다. 그러므로 個人에 붙여지는 코우드 등이 이 例인데 여기에서의 코우드는 반드시 特定한 個人을 意味하며 그 외의 어떤 다른 사람도 나타내지 않는다. 이 코우드에는 將來 變化될 可能性이 있는 要素를 可能な 限 包含시키지 않는 것이 좋다.

● 分類를 위한 코우드

어떤 体系 혹은 方法에 따라 데이터 項目을 種類別로 分類, 區分하기 위하여 設定된 코우드로 데이터의 体系化를 위하여 많이 使用되고 있다. 자리數別 코우드는 이것의 典型的인 例이다, 여기에서 注意할 點은 比較的 上位의 分類階層은 例外이나, 대체로 用途나 目的에 따라 變動이 많으므로 미리 그러한 것을 明確히 해 두어야 할 일이다. 또한 彈力性이 있는 코우드 体系를 만드는 하나의 方法은 比較的 上位의 階層을 標準化하고, 下位의 階層은 各各의 部門마다 그 機能에 따라 適切하게 코우드 体系를 設定하는 것이다. 또한 意味없는 分類를 위한 分類의 階層에 빠지는 일이 없어야 할 것이다.

● 排列을 위한 코우드

一定한 順序로 들어 놓거나 階級이나 時間에 따라 排列하기 위한 코우드로 가나다順序나 發生順序를 나타내는 코우드 등이 이에 속한다. 이 코우드에도 變化하지 않는 것과 變化 可能性이 있는 것이 있기 때문에 後者의 경우 追加, 挿入 등에 의하여 코우드의 体系가 무너지지 않도록 分類코우드에서와 같은 注意가 必要하다.

3.1 데이터 要素의 코우드化

앞에서는 將來 變換可能性이 있는 要素를 可能な 限 코우드에 包含시키지 않을 것을 強調하였는데, 이것을 최소한 防止하고 또한 從來의 統計的, 集計的인 데이터 코우드에서 脫皮하고 오늘날의 온라인·리얼타임 및 타임 셰어링의 普及과 빠른 速度의 大容量記憶裝置에 따른 데이터 뱅크의 實用化와 함께 데이터의 共同利用 및 情報檢索 등의 새로운 시스템에 適用될 수 있는 데이터 코우드의 体系가 開發되고 標準化되어야 한다. 예를 들면 온라인·리얼타임에서, 멀리 떨어진 곳으로부터 中央의 데이터 뱅크에 要求가 있는 경우에는 數秒以內에 여러 가지 變動이 많은 데이터 項目 要素에서 그 당시에 不充分한 情報를 데이터 뱅크의 파일과 對照하여 要求된 情報를 즉시 찾아 회답해 주어야만 하는 일이다. 이와 같은 경우 識別의 要因이 되는 데이터 項目의 要素를 데이터 要素라 부른다. 데이터 項目이란 個人의 例를 들어 說明하면, 어떤 特定個人이 他人과 구별되는 속성으로 姓名, 生年月日, 住所, 本籍地, 職業, 經濟力, 結婚關係, 血液型, 性別, 指紋, 身長, 体重 등인데 이中 個人의 識別을 위해 共通으로 使用할 수 있는 것은 個人識別의 데이터 要素가 된다.

데이터 要素의 條件으로는 다음과 같은 것들이 있다.

- (1) 상당한 기간동안 변하지 않는 것으로 생각되는 것
- (2) 최초 파일 設定때에 確實한 것이, 入手될 수 있는 것
- (3) 파일로부터 쉽게 모순되는 것을 뽑아낼 수 있는 것
- (4) 單一의 데이터 要素에서는 무리이나 여러 개

의 데이터 要素를 組合하였을 때 特定한 데이터 項目을 識別할 수 있는 것

이와 같은 條件으로 선택된 데이터 要素를 코우드化할 때에는 다음 페이지에 있는 表 2에 표시된 것과 같이 예전부터 適用되어 오던 데이터 코우드의 형태를 그대로 適用할 수 있다.

4. 情報處理 標準化의 國際的인 動向

4.1 國際的인 現況

國際標準化機構 (ISO)의 第97技術委員會 (TC 90)는 Computers and Information Processing 을 맡아서 標準化를 推進하고 있는데 TC90은 14개의 小分科委를 두고 있다. 여기에서 制定된 國際規格은 現在까지 56件이다.

- 國際電氣標準會議 (IEC)는 電子計算機의 하드웨어에 關係되는 部分의 標準化를 推進하고 있다.

- 國際通信聯盟 (ITU)은 데이터傳送에 關係한 標準化를 위하여 努力하고 있다.

- 유럽컴퓨터工業會 (ECMA)는 電子計算機 및 情報處理의 標準化를 위한 研究開發에 從事하고 있으며, 여기서 研究된 結果가 ISO에 의하여 採擇되는 경우 ISO規格이 된다.

- 國際情報處理聯盟 (IFIP)은 學術活動이 주된 業務지만 標準用語 및 프로그래밍 言語 등에 關係한 標準化 作業結果를 ISO에 提案하여 採擇된 바 있다.

4.2 美國의 現況

美國國家規格協會 (ANSI)의 X3分科委員會가 研究制定한 規格이 一般的으로 國家規格으로 通用되고 있으며, 이것이 美標準局 (NBS)에 의하여 採擇되는 경우는 聯邦規格이 된다.

現在까지 36件의 ANSI規格이 制定되었다.

- ANSI規格

이 規格은 強制力이 없으나 NBS에 의하여 採擇되는 경우에는 政府機關에 대하여 強制力을 行할 수 있도록 되어 있다. 이러한 法的인 바탕

으로는 이른바 브룩스法 (Brooks Bill)이 있는데 이 法律은 1965年 10월에 公布되었으며, 그 內容은 “政府 各機關이 個別的으로 調達해 온 컴퓨터를 一括 契約토록 하였으며, 컴퓨터 및 情報處理關係의 規格開發에 關係한 權限賦與”에 關係한 것이다.

4.3 日本의 現況

日本情報處理學會, 情報處理開發센터, 日本電子工業振興協會, 日本소프트웨어協會 등에 研究委託하여 作成된 原案을 日本工業標準調查會가 檢討한 후 이를 主務官廳을 通하여 日本規格協會가 JIS로 公示한다.

4.4 其他

英國은 英國標準研究所 (British Standard Institution)에서 BS로서, 獨逸은 DIN規格으로, 프랑스는 AFNDR, 스웨덴은 中央統計局 등에서 데이터 코우드의 統一化를 試圖하고 있다.

5. 우리나라의 現況

우리나라에서의 情報處理標準化, 특히 데이터 코우드 統一化에 關聯된 機關으로는 科學技術處 情報產業局, 總務處 行政電算擔當, 經濟企劃院 統計局, 商工部 工業振興 標準擔當 등이 있으나 各 部署에서의 統一化를 위한 接近方式은 機能에 따라 다르다.

그러므로 情報處理 標準化는 모든 사람에게 의해 그 必要性이 力說되고 있으나 뚜렷한 解決方案은 나오지 않고 있다. 또 나온다고 하더라도 위의 各 部署의 機能에 多少의 重複은 반드시 생기게 마련이다. 가령, 物品코우드의 統一化라는 事業이 있다면 이것은 어느 部署가 主役이 되어야 하는가에 대하여 뚜렷한 定義를 내리기가 힘들다.

그러므로 이러한 事業이 一元化되면서 強力히 推進할 수 있는 制度的인 面이 講究되어야 할 것이다.

한편 國內政府機關의 코우드 使用現況은 다음 表 3과 같다.

表 2. 데이터 코우드의 形態와 特徵

形 態	說 明	特 徵	適 用 分 野
順序 코우드	發生順, 크기順, 알파벳順 등의 一定한 基準에 따라 배열한 순위로 記號 혹은 숫자의 일련번호를 붙인다	1. 簡單, 明瞭 2. 자리수가 적다 3. 發生順일 경우에는 추가가 편리 4. 識別이 容易 5. 分類에 부적합	1. 分類基準이 確立된 곳 2. 다른 코우드와 組合되어 사용되는 일이 많다
區分 코우드	一連番號를 미리 끊어서 共通의 特性을 가진 데이터 項目群을 그 부분에 넣어 준다	1. 적은 자리수로 많은 그룹을 나눌 수 있다	1. 자리수에 制限이 있을 때 그룹으로 나누는 경우 2. 다른 코우드와 組合되어 사용되는 일이 많다
자리수별 코우드	데이터의 項目을 어떤 정해진 基準에 따라 分類하여 分類의 階層을 자리수로 표시하고 各各의 그룹내에서 連續番號를 붙인다	1. 構成分類基準이 明確하다 2. 理解하기 쉽다 3. 기계처리에 가장 적합 4. 追加가 容易 5. 자리수가 많아지기 쉽다	1. 分類基準이 明確한 때 가장 利用度가 높다
10進 코우드	10進數의 각 자리수에 分類의 階層을 주고 下位 계층의 追加를 무한으로 확대할 수 있도록 한 것으로 圖書의 整理 및 주소의 番地와 같은 것에 붙여지고 있다	1. 追加가 容易 2. 무한으로 분류의 계층을 細分化 할 수 있다 3. 자리수가 많아진다	1. 圖書整理, 住居表示, 파일의 정리에 사용된다 2. 데이터 項目의 情報가 조밀할 때
表의 코우드	데이터 項目의 속성을 직접 혹은 간접으로 표시하는 숫자, 문자, 記號를 組合한 것	1. 內容이 표시하고 있는 것을 쉽게 解讀할 수 있다 2. 자리수가 많아지기 쉽다	1. 性能, 單位, 치수, 重量, 距離 등에 사용된다
數字式 文字 코우드	一定한 방식에 따라 문자를 숫자로 바꾸어 놓은 것	1. 알파벳, 가나다順의 索引이 容易 2. 자리수가 많다	1. 人名, 法人名 등에 사용된다
合成 코우드	어떤 코우드의 前後에 다른 코우드를 붙여 合成하는 것으로 데이터 項目을 여러 각도로 識別하고 處理할 수 있도록 한 것	1. 分類의 階層을 增加시켜 多角的으로 識別할 수 있다. 2. 자리수가 많아지기 쉽다	1. 多角的으로 固有性이 있는 것에 利用된다

表 3. 機關別 코드 使用現況

機關別 / 區分	種 類	制定機關	基 準	体 係
經濟企劃院	<ul style="list-style-type: none"> ● 標準行政區域分類 ● 韓國標準產業分類 ● 韓國標準職業分類 ● 韓國標準商品分類 ● 韓國標準疾病死因分類 ● 國際貿易分類 	自 體 “ “ “ “ 財 務 部	全國行政區域 UN 國際產業分類 UN 國際職業分類 SITC, BTN	Serial Code “ “ “ “
外 務 部	<ul style="list-style-type: none"> ● 國家, 旅行目的分類 ● 學歷, 職業分類 ● 在外公館分類 	自 體 “ “	業務基準 “ “	BLOCK 年 別
治 安 本 部	<ul style="list-style-type: none"> ● 地域, 官署分類 ● 罪名, 手法分類 	自 體 “	法定洞 Code 檢察廳 “	
財 務 部	<ul style="list-style-type: none"> ● 會計所管別分類 ● 豫算別分類 ● 會計職公務員分類 ● 物品分類 	經濟企劃院 “ 自 體 “	FSC 美聯邦補給分類	BLOCK “ “
國 防 部	<ul style="list-style-type: none"> ● 地域分類 	內 務 部	法定洞 Code	BLOCK
農 水 產 部	<ul style="list-style-type: none"> ● 市, 道別分類 ● 農家別分類 	自 體 “		Serial 大·中·小分類
專 賣 廳	<ul style="list-style-type: none"> ● 計定番號, 官署別分類 ● 等級, 棄分, 年產分類 ● 事業, 原價, 機械分類 ● 物品分類番號 ● 담배製品分類番號 ● 原料 및 담배分類 	自 體 “ “ 財 務 部 自 體 “	政府物品分類基準	BLOCK Code “ “ Serial “ “
國 稅 廳	<ul style="list-style-type: none"> ● 納稅者番號分類 ● 商品品目分類 ● 家口分類 ● 市, 郡分類 ● 稅關, 財源 및 形態別分類 ● 關稅 및 商品別分類 ● 學歷, 職業別分類 ● 犯罪類型, 檢舉場所分類 ● 稅關職員別分類 ● 其 他 	自 體 “ “ “ 自 體 財 務 部 自 體 “	美國社會保障制度分類 所得標準率分類 대만가구 Code 國內地域 SITC, BTN	Serial BLOCK “ “ Serial “ “ “
勞 動 廳	<ul style="list-style-type: none"> ● 職種, 產業, 番號別分類 ● 地域別分類 ● 學科, 免許, 個人別分類 	經濟企劃院 內 務 部 自 體	UN 標準國際分類 法定洞 “	Serial “
鐵 道 廳	<ul style="list-style-type: none"> ● 全國驛名分類 	自 體	距離別基準	Serial

	●貨物品目, 貨車, 機關車 分類 ●物品分類 ●個人別分類	“ 財 務 部 自 體	용적별基準 F S C 人事基準	BLOCK Serial BLOCK
서울市	●行政洞別分類 ●特定洞別分類 ●自動車登録番號分類 ●家屋構造, 用度別分類 ●水道業種別分類	自 體 “ “ “ “	地方稅法基準 “ “ 國稅廳, 不動產課稅 標準額 上水道給水條例	BLOCK “ “ “ Serial

6. 時急히 코우드化 되어야 할 데이터

우리나라에서의 사람에 관련된 코우드는 住民番號가 制定되어 活用段階에 이르렀고 이것은 生年月日, 男女區分, 地域코우드에 관한 情報를 內包하고 있으며 이것은 一般的으로 가장 널리 쓰이는 사람의 識別方法으로 通用되기 시작했다. 그러나 그밖의 追加的인 情報는 이것만으로는 알 수 없으므로 경우에 따라서는 다른 意味의 코우드를 追加해서 使用하여야 한다. 이러한 事例는 사람에 관해서만 適用되는 것이 아니다.

다음은 우리나라에서 時急히 統一코우드를 制定하여야 할 데이터를 나열하여 보았다. 이러한 基本的인 데이터를 必要에 따라 몇 개씩 組合使用할 수 있도록 定하는 것이 바람직하다고 하겠다.

(1) 사람에 關連되는 코우드

주민번호, 국적, 주거표시, 혈액형, 장애등급, 세대, 가족관계, 학력, 직업, 혼인, 성격, 인사급여, 직종, 직급, 기능, 면허자격 등

(2) 물품에 關連되는 코우드

재료, 상품, 수송기종

(3) 경제통제에 關連되는 코우드

산업분류, 기업체, 예산회계과목, 기업형태, 고정자산품목, 거래결재조건 등

(4) 토지건물에 關連되는 코우드

지목, 용도, 이용, 획지조건, 건물의 조건, 건물의 용도, 권리의 종류, 지역메슈 등

(5) 기타 공통적으로 적용되는 코우드

시간, 날짜, 날씨, 국명, 관공서, 학교교육기관, 은행금융기관, 병원, 국제기관 등

7. 結言

우리나라에서의 데이터 코우드 標準化 事業은 一次的으로 政府가 先導하는 立場에서 政府 各機關에서 統一시키고 이를 民間機關에서도 適用하도록 勸奨한 다음, 國家規格화한다는 方針아래서 推進되어야 할 것으로 생각된다. 國家規格이 先行되어도 이것이 充分히 適用되지 않는 경우가 많다. 例를 들면 이미 蓄積段階에 있는 데이터 파일(Data File)의 各種 데이터 要素를 統一規格대로 訂正하는데 있어서 所要되는 人力, 經費 및 時間이 經濟的이 되지 못한 경우이다. 따라서 統一化 作業에 所要되는 經費에 대한 合理的인 措置가 뒤따르지 못하는 경우에는 統一化 作業은 失敗하고 만다.

一般的으로 統一化 및 標準化 過程은 메이커와 利用者 間의 利害相反에 의한 싸움의 連續이다. 統一 및 標準은 이미 利用되고 있는 다른 方法을 無視해 버리는 경우가 있기 때문에 어려운 것이다.

이제 걸음마를 시작한 國內의 情報産業과 컴퓨터 利用者 集團은 現在 努力을 기울이고 있는 데이터파일 構成에 앞서 統一化에 積極 協助하여야 할 것이다.

끝으로 政府는 이를 위한 制度를 確立함과 아울러 統一에 隨伴되는 所要經費의 補償策도 同時에 講究함으로써 이 事業이 成功할 수 있을 것임을 부언한다.

參 考 文 獻

과학기술처, EDPS 표준화사업을 위한 조사연구, R-75-74, 1975.