

商用 데이터베이스: 그 要點과 活用(1) — 데이터베이스 서비스의 動向*

Commercial Databases : The Keypoints and Practical Use(1)
— Trends of Database Services —

趙 載 浩 譯**
(Cho, Jae Ho)

抄 錄

우리는 情報에 근거를 두고 행동을 취하고 있다. 따라서 효율적인 행동을 하려면 정보의 入手 루트를 잘 가꾸어야 한다. 여러 經路의 루트가 있지만, 특히 商用 데이터베이스는 利點이 크다. 더구나 대가를 지불하면 누구든 자유롭게 사용할 수 있으므로, 이용 여하에 따라서 情報의 입수에 상당한 隔差가 생기게 된다. 한편, 日本의 商用 데이터베이스 서비스는 출현된지 20여년이 지나 최근 많이 변화되었다. 예를 들면, 배치에서 온라인으로 그리고 CD-ROM의 등장에 의한 미디어의 多樣化는 그 대표적인 예이다. 데이터베이스의 최신 동향을 살펴, 情報化時代의 틀(道具)으로써 活用하는 것이 매우 중요하다.

키 워 드

데이터베이스, 情報利用, 情報蒐集, 情報 서비스, 主題知識

ABSTRACT

Nowadays we usually act based on information in order to act effectively, we should enlarge our access to information of course, we have several methods to choose from, but commercial databases are the most advantageous. One reason is that everyone has equal access to commercial databases. As a result, people or companies who use commercial database information skillfully have a distinct advantage over those who don't commercial database ser-

* 本稿는 日本科學技術情報센터(JICST)에서 발간하는 「情報管理」, 第36卷 1號, pp. 49~60(1993. 4. 1)에 수록된 日本情報處理開發協會 鈴木 茂樹씨의 “データベース・サービスの動向”을 번역한 것으로 데이터베이스의 이해에 도움이 되고자 앞으로 12회에 걸쳐 연재될 예정이다.

** 産業技術情報院, 情報資料室 招請專門委員.

vices in japan have been changing rapidly since their birth more than twenty years ago. For example, the distribution mode of databases is changing : from the original batch(off line) to on-line and most recently to the coexistence of batch and on-line with the advent of CD-ROM. It is important for us to keep up with the new trends in databases, and to utilize them effectively as a tool in the information age.

KEYWORDS

Database, Infomation Use, Information Collection, Information Service, Subject Knowledge.

I. 序 論

1. 데이터베이스 環境의 變化

이번부터 12회에 걸쳐 「商用 데이터베이스 : 그 要點과 活用」을 연재케 되었다. 本 講座에서는 7년 전에 「데이터베이스」를 다루었으나, 그 동안에 데이터베이스를 에워싼 환경은 크게 변화되었다. 그와 관련해서 通産省의 데이터베이스 臺帳에 의하면 이용이 가능한 데이터베이스 數는 1,289(1985년도)種에서 2,686(1991년도)種으로 2배 이상이 되었다. 특히, 日本의 데이터베이스는 同 期間에 281種에서 892種으로 3배 이상 増加하였다.

또한 데이터베이스의 提供形態도 크게 변화되었다. 데이터베이스 서비스는 처음에는 磁氣 테이프에 의한 배치(오프라인) 제공으로 출발하였다. 이어서 通信回線을 통한 온라인이 主流가 되었으나, 최근에는 CD-ROM의 출현으로 온라인과 오프라인의 共存時代를 맞이하게 되었다. 더구나 퍼스컴 通信의 보급으로 데이터베이스 檢索이 더욱 이용자와 가깝게 되었다.

데이터베이스의 종수가 증가하여 提供形態가 다양화되고, 가까운 것이 되면 될수록 그 이용 여하에 따라서 情報의 입수에 차이가 생기게 된다. 이번의 강좌에서는 데이터베이스 環境의 變化를 고려하면서 수요가 큰 데이터베이스를 분야마다 다루어 데이터베이스의 活用의 要點을 소개하고자 한다.

2. 情報入手의 루트

우리의 주위에는 生活現場이나 비즈니스 현장을 불문하고 다양한 데이터가

넘치고 있다. 그러나 모든 데이터가 모든 사람에게 관계를 갖고 있는 것은 아니다. 홍수처럼 넘치는 ‘데이터’ 중에서 그때 마다의 행동에 필요한 것을 입수하여 ‘情報’로 轉化하는 것이 중요하다.

즉, 데이터와 情報는 다르다. 데이터는 어디에도 있으나, 情報는 그렇지 않다. 데이터 중에서 自己(end user)에게 필요한 것, 그리고 어떤 행동에 맺어지는 것, 그것이 情報이다. 더구나 데이터를 정보로 轉化하는 작업은 이용자 자신이 하지 않으면 안된다.

이러한 데이터를 大別하면, ①自然現象, ②文獻/메스 미디어, ③人脈, 그리고 데이터베이스 중에 있다. 예를 들면, 아침에 일어나 하늘의 상태를 보고 우산을 갖고 외출하는 행동이 되면 그 사람은 모든 사람에게 공통된 하늘의 상태라고 하는 ‘데이터’로부터 ‘오늘은 비가 올 것같으니 우산을 갖고 간다’라는 ‘정보’를 입수한 셈이 된다.

文獻이나 메스컴도 萬民共通의 데이터로서 몸 가까이 넘치고 있다. 그러나 圖書館에 있는 모든 문헌이 우리에게 필요한 것은 아니다. 이처럼 모든 新聞記事 또는 TV 프로가 불가결한 것이 아니다. 利用者의 취미나 비즈니스에서 추구하고 있는 테마에 따라서 필요한 것이 한정되는 것이다. 오히려 넘치는 데이터 중에서 어떻게 하여 필요한 정보를 뽑아 내느냐가 중요하다. 예컨대 文獻에서는 자기에게 도움이 되는 「코어 저널」을 갖는 것도 중요하다.

人脈은 정보입수에 있어서 크게 도움이 된다. 人脈의 한사람 한사람은 그 어떤 전문가이다. 들으면 곧바로 응하여 준다. 다만 人脈은 하루 아침에 이루어지는 것이 아니다. ‘Give and Take’는 원래부터 相應의 노력과 시간이 소요된다. 그런데, ‘데이터베이스’는 이름 그대로 정보는 아니고, 데이터의 寶庫이다. 自然現象의 관측 結果는 본래부터 文獻, 新聞記事, 무엇이든 채워져 있다. 나아가서 換이나 株價 등 시시각각으로 변하는 데이터도 들어가 있다.

데이터베이스는 ①個人이 자기를 위하여 作成하는 個人 데이터베이스, ②特定の 기업 등이 交叉로 이용되는 인 하우스 데이터베이스, ③비용을 지불하면 누구든 이용되는 ‘商用 데이터베이스’가 있다.

특히 商用 데이터베이스는 모든 利用者에게 개방되어 있어 이의 활용 여하에 따라서 情報의 입수에 상당한 차이가 생긴다. 현재로선 端末機 하나만 있으면 세계의 온갖 商用 데이터베이스가 온라인으로 이용되는 상황이므로 더욱 差異가 생기게 마련이다.

II. 데이터베이스란?

데이터베이스라는 말은 1950년대에 美國防省에서 사용되기 시작한 뒤, 일반화되었다. 즉, 그 당시 세계에 전개되고 있던 兵器나 武器 등에 관한 情報를 집중 관리하기 위하여 國防省은 컴퓨터를 驅使한 라이브러리를 개발하였다. 이것을 데이터의 基地(베이스)라는 의미로 데이터베이스라고 부르게 되었다.

日本の 著作權法에는 다음과 같이 정의하고 있다. 「論文, 數值, 圖形 기타의 情報의 集合體로서 그것들의 정보를 電子計算機를 이용하여 檢索될 수 있도록 체계적으로 구성한 것」

이러한 語源이나 定義에서는 컴퓨터가 기본이 되어 있음을 알 수 있으나, 데이터베이스의 정확한 이미지는 분명치 않다. 이에 데이터베이스를 몇개의 觀點에서 分類해 보면, 데이터베이스란 무엇인지 明確해진다.

1. 데이터베이스의 分類

데이터베이스는 그 형태나 用途에 따라서 몇 개의 類型으로 분류할 수 있다. 데이터베이스는 그 형태에 의해, ①레퍼런스 데이터베이스 ②팩트 데이터베이스로 대별된다.

(1) 레퍼런스 데이터베이스

記事나 論文 등 문헌의 書誌事項이나 抄錄 등을 수록한 것이다. 즉, 오리지널 文獻의 안내 정보를 제공하는 것이다. 이런 뜻으로 文獻 데이터베이스라 말하기도 한다.

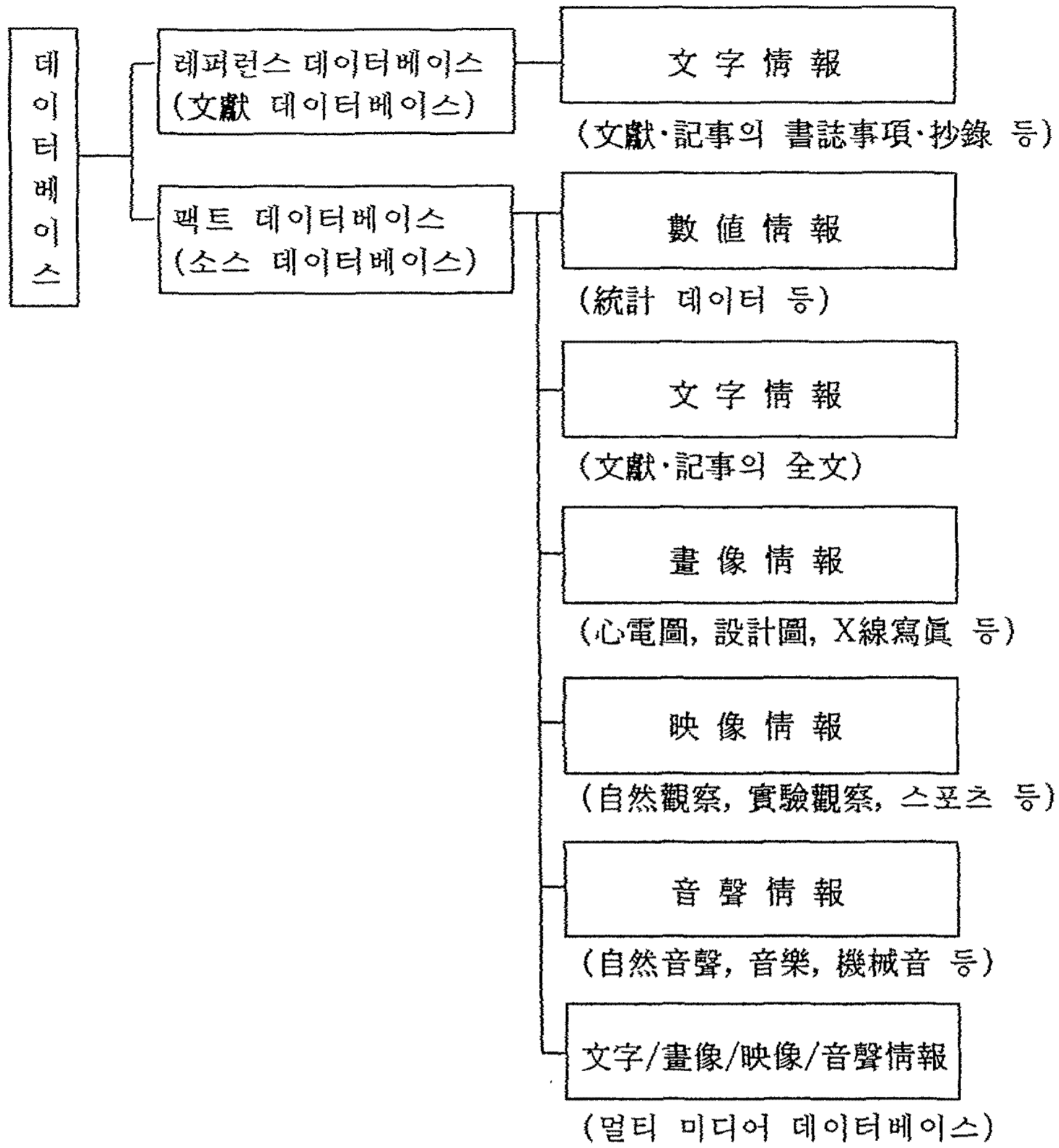
팩트 데이터베이스는 오리지널 情報 그 자체를 제공하는 것이다. 즉, 文獻의 경우는 全文을 수록하고, 그 밖에 統計 등의 數值나 畫像, 映像, 音聲 등을 저장한 것도 있다. 또한 이것들을 혼합한 것은 「멀티 미디어 데이터베이스」로서 注目을 받고 있다(〈圖 1〉 참조).

(2) 提供 모드별 分類

데이터베이스의 提供形態에 착안한 분류로 온라인과 오프라인으로 대별된다. 온라인 데이터베이스는 데이터베이스 호스트가 通信回線을 통하여 端末機와 접속되고 있는 형태이다. 한편으로 오프라인 데이터베이스는 自立型이라고

〈圖 1〉

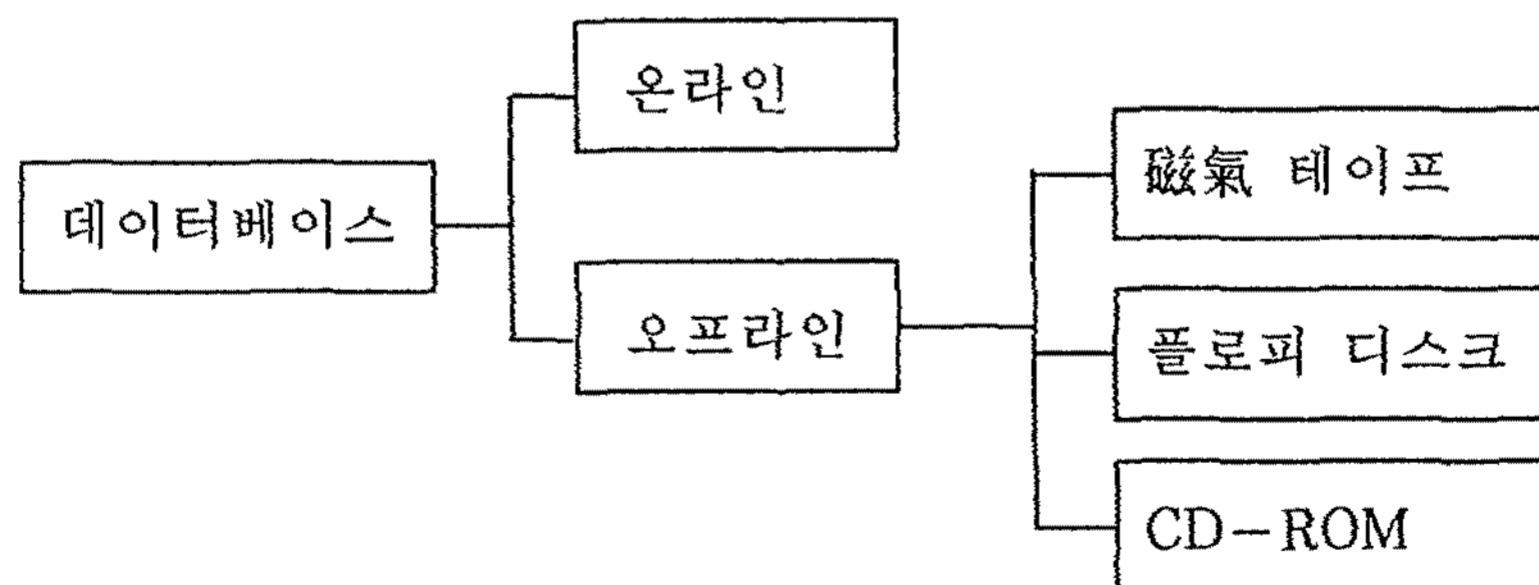
데이터베이스의 分類



資料 : 데이터베이스 振興 센터의 멀티 미디어 데이터베이스를 根據로 작성.

〈圖 2〉

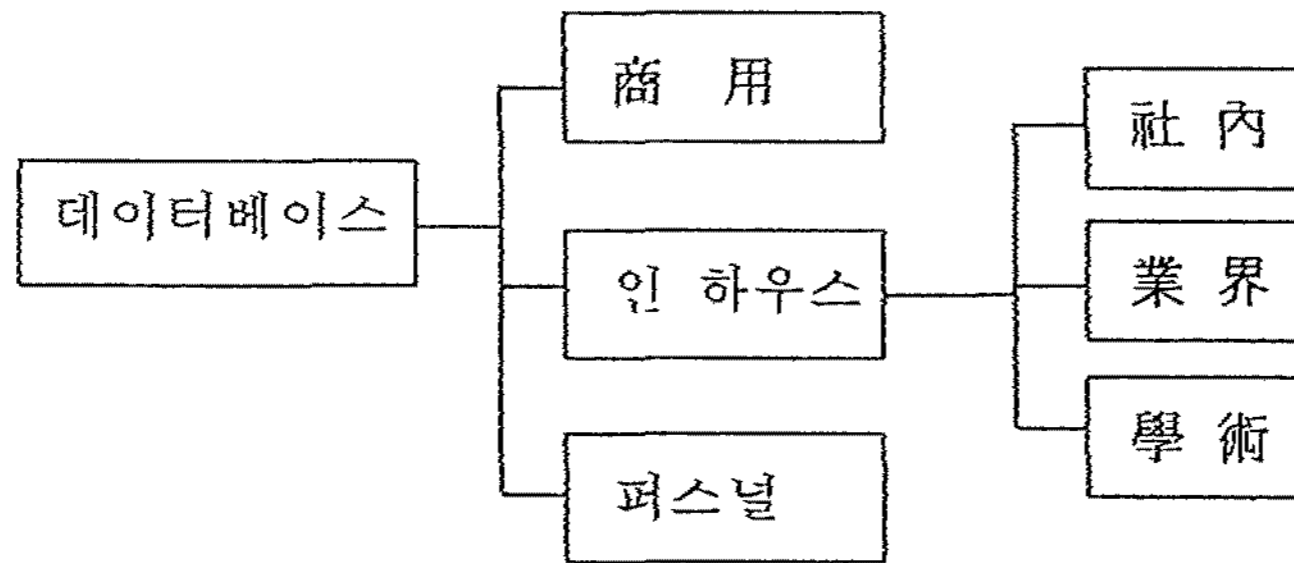
提供形態別 分類



부르며, 데이터베이스가 磁氣 테이프(MT), 플로피 디스크(FD), CD-ROM 등의 媒體에 저장되고 있다(〈圖 2〉 참조).

〈圖 3〉

用途別 분류



(3) 用途別 분류

데이터베이스가 어떠한 목적으로 構築되어 이용되고 있는가, 즉 데이터베이스의 용도에 착안한 분류로 크게 ①商用 데이터베이스, ②인 하우스 데이터베이스, ③퍼스널 데이터베이스로 나눈다.

이 講座에서 다루는 商用 데이터베이스는 모든 이용자에게 開放되고, 대가를 지불하면 누구든 이용할 수 있다. 인 하우스 데이터베이스는 企業內 데이터베이스, 學術 데이터베이스, 官廳 데이터베이스, 業界 데이터베이스 등 몇 개의 종류가 있다. 어느 경우든 모든 이용자에게 개방된 것이 아니고 이용자는 한정된다.

퍼스널 데이터베이스는 문자 그대로, 個人用의 데이터베이스이다. 이용자 자신이 이용하기 때문에 퍼스컴 등을 사용하여 스스로가 作成하는 데이터베이스이다. 規模는 적지만 자신의 니드에 흡사한 데이터만이 저장되어 있다. 이런 뜻으로 퍼스널 데이터베이스는 ‘데이터’보다는 오히려 ‘情報 그 자체’가 들어가 있다고도 말할 수 있다(〈圖 3〉 참조).

(4) 分野別 분류

데이터베이스에 저장되어 있는 情報의 分野에 따른 분류이다. 예를 들면, 科學情報를 다루고 있으면 科學技術 데이터베이스, 비즈니스 情報이면 비즈니스 데이터베이스라고 부른다. 또는 비즈니스 중에서도 특히 株價나 換 등에 특화된 것은 金融 데이터베이스라 불리울 때도 있다.

데이터베이스의 내용이나 特色을 소개하고 있는 디렉토리는 이 분야별 분류가 중요하다. 그와 관련해서 美國의 Cuadra/Gale이 발행한 *Directory of On-Line Database*에는 주제별로 자세한 분류를 채용하고 있다.

日本에서는 通産省이 매년 발행하고 있는 「데이터베이스 臺帳總覽」이 대표적인 디렉토리이다. 同 臺帳에는 데이터베이스를 「大分野」와 「細分野」의 두

분야로 분류하였다. 전자에는 ①一般, ②自然科學·技術, ③社會·人文科學, ④비즈니스 등 4분야가 있고, 후자인 細分野에는 主題에 따라서 도합 58개의 분야가 마련되어 있다.

2. 데이터베이스의 機能과 特色

데이터베이스는 컴퓨터의 기능을 最大限으로 驅使하여 이용자에게 여러 가지의 便益을 제공하고 있다. 데이터베이스의 주요한 특색에는 網羅性, 多面檢索性, 遠隔操作性 등이 있다.

(1) 網 羅 性

문헌의 예를 들면 과거의 오랜 것에서부터 최신의 것까지, 또한 동시에 세계의 많은 종류의 것을 커버할 수 있는 기능이다. 大容量의 記憶 미디어를 기본으로 하고 있다.

현재처럼 多種多樣한 문헌이 나오고 있는 상황에 있어선 개개 이용자가 모든 문헌을 추적한다는 것은 사람의 限界를 넘어선 것이다. 물론, 한 사람이 그 만큼의 데이터베이스를 構築한다는 것도 불가능하다. 결국 網羅性이라는 기능에 의하여 많은 사람들이 長期間에 걸쳐 작성한 데이터베이스를 마치 자기 것처럼 이용할 수 있게 된다.

(2) 多面 檢索性

데이터베이스에 저장되고 있는 대량의 데이터 중에서 자기에게 필요한 情報를 뽑아내기 위한 기능으로 컴퓨터의 計算機能이 기본이다.

데이터베이스는 網羅性이 있으므로 단순한 檢索에도 대량의 데이터가 出力되어 버린다. 이를 피하기 위해선 키워드의 적절한 結合으로 효과적인 검색이 불가결하다. 이러한 접근 수법으로 참으로 필요한 정보를 複數의 관점에서 입수할 수 있게 된다.

(3) 遠隔 操作性

데이터베이스가 저장되어 있는 호스트 컴퓨터가 어디에 있더라도 가까운 端末機로부터 정보가 입수되는 기능이다. 컴퓨터와 電氣通信網이 일체가 된 온라인 處理機能이 기본이다.

이 기능에 의하여 데이터베이스가 해외에 있더라도 必要時에는 자유롭게 이용할 수 있다. 이상의 特色이 동시에 채워졌을 때, 이용자는 세계의 정보를 자기의 손 안에 쥐고 있는 기분이 들기도 한다.

3. 廣義의 데이터베이스

데이터베이스를 에워싼 環境은 극히 流動的이다. 예컨대, 얼마 전만 해도 데이터베이스라면 磁氣 테이프에 저장된 소위 배치 베이스의 것이 었으나, 지금은 온라인 데이터베이스가 主流를 이룬다. 나아가 최근에는 CD-ROM의 등장으로 인해 自立形態(오프라인)에의 이용도 伸張되고 있다. 즉, 온라인과 배치의 공존시대에 들어선 셈이다.

또한, 온라인 리얼 타임이 주류라고 말하여도, 그 수단은 電氣通信 回線에 한정된다. 예컨대, 무선을 사용한 株價情報 서비스 등은 이미 제공되고 있다. 더구나 TV의 放送電波에 실려 텔레텍스트(文字 서비스)에 의한 각종 정보 서비스도 출현하고 있다. 앞으로는 衛星이나 CATV 등에 의한 데이터베이스 서비스도 출현될 것이다.

이와 같이 배치 베이스에도 CD-ROM 뿐만 아니라, IC 카드에 의한 개인의 의료 데이터베이스 등이 登場하고 있다.

또한 퍼스컴 通信 등 가까운 수단을 통하여 데이터베이스에 接續하는 사람들도 증가되고 있다. 더구나 어느 데이터베이스 호스트와 계약하면, 다른 호스트의 데이터베이스도 檢索하여 이용되는 게이트웨이 서비스의 인기도 높아지고 있다. 프랑스에서는 데이터베이스 서비스라면 텔레텔이라 일컫는 비디오텍스가 주류를 이루고 있다.

데이터베이스에 수용되는 데이터(情報)도 변화되고 있다. 人工知能(AI) 技術을 구사하여 專門家の 知識(노하우)을 저장한 「知識 베이스」같은 것도 실현되고 있다. 動畫나 음성 등의 데이터베이스, 혹은 이것들과 기존의 데이터베이스를 혼합한 「멀티 미디어 데이터베이스」도 本格化될 것이다. 人間の 苦惱(感性)의 기능을 갖춘 기술의 개발도 시작되고 있어, 장래에는 「感性 데이터베이스」라는 것도 등장될 가능성도 있다. 이처럼 데이터베이스를 에워싼 環境은 앞으로 더욱 더 변화될 가능성이 높다.

위와 같은 流動的 상황을 고려한다면 「데이터베이스란 무엇인가?」에 대하여 限定的인 테를 끼우기보다는 廣義로 생각하는 편이 좋지 않을까 한다. 여기서는 基本的으로 아래의 두 가지 점을 동시에 充足시키면 데이터베이스라 하기로 하자. 즉, ①데이터(또는 정보)가 電子化 되어있는 것, ②이용자의 어떤 행위에 의해 필요한 부분을 효율적으로 끌어낼 수 있는 것.

다만, ‘동시에 충족하는’것이 중요하고, 단지 電子化 되어 있다고 해서 데이

터베이스라고 할 수는 없다.

Ⅲ. 데이터베이스 産業의 動向

1. 發展 經緯

日本の 데이터베이스 서비스는 1970년에 시작되었다. 이 해에 日本經濟新聞社가 株式 데이터 등을 磁氣 테이프에 판매하기 시작하였다. 日本科學技術情報센터(JICST)는 1957년에 科學技術情報의 자기 테이프 제공을 시작하였다. 이 해에 特許情報機構도 特許情報의 배치 서비스를 개시하였다. 이 해가 일본의 데이터베이스 서비스의 여명기라 하겠다. 거의 같은 시기에 온라인 서비스도 시작되었다. 즉, 1972년에는 日經의 NEEDSTS(企業財務情報)를 효시로 1973년에는 전통의 MARK II(海外情報), 1974년엔 QUICK의 비디오 I(株價情報), 1976년에는 JICST의 JOIS-I 등이 판매되었다.

1979년에는 民間 데이터베이스 업자의 단체로서 데이터베이스 서비스업 連絡懇談會(현 日本 데이터베이스協會/DINSA)가 通産省의 주선으로 설립되었다. 1980년대에 들어가 데이터베이스를 에워싼 환경은 갑자기 활기를 띠기 시작하였다. 그와 관련하여 通産省은 産業構造審議會의 소위 「1980년대 비전 答申」에서 데이터베이스의 중요성을 나타냈다.

나아가 1983년에는 日本에서 이용되는 데이터베이스의 편람으로서, 「데이터베이스 臺帳總覽」을 간행하였다.

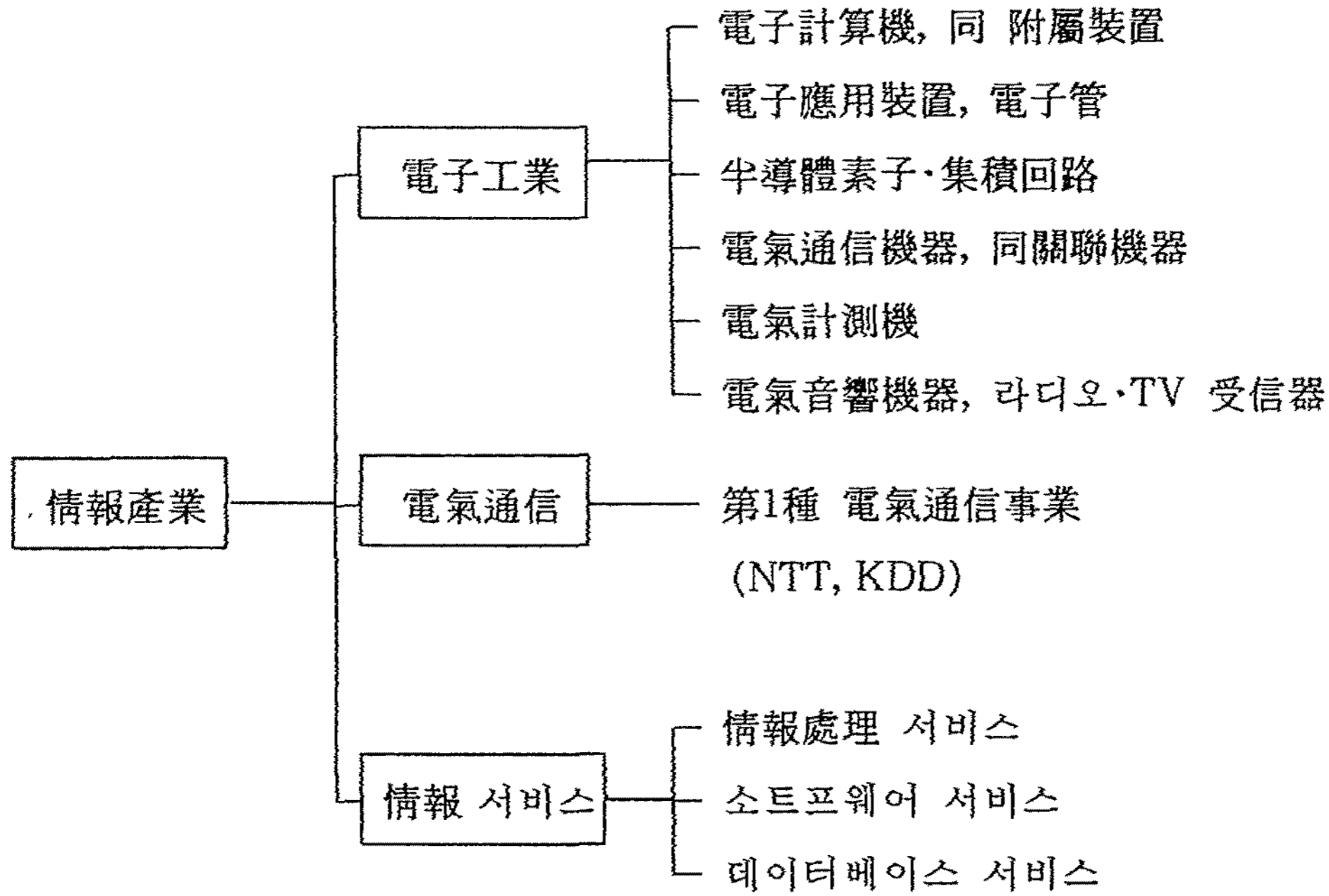
이러한 일련의 활동으로 인해, 1980년대 중반에 이르러 日本의 데이터베이스 서비스는 본격적인 시기를 맞이하였다. 이 시기에는 상당한 기업이 대거 데이터베이스에 참여하였다.

2. 데이터베이스 産業의 位置와 賣出額

데이터베이스 産業은 「情報産業」의 일익을 담당하는 「情報 서비스업」의 한 업종으로서 위치가 부각되었다. 정보산업의 범위에 관해선 몇 가지 다루는 방법이 있으나, 産業構造審議會에는 ①電子工業, ②電氣通信業, ③情報 서비스업을 情報産業의 세 기둥으로 하고 있다(〈圖 4〉 참조). 그 밖에 新聞·放送 등

〈圖 4〉

情報産業의 분류



資料:通産省 産業構造審議會.

의 매스컴이나 雜誌·書籍 등의 출판을 포함하는 경우도 있다.

日本の 데이터베이스 賣出額은 通産省의 지정통계 「特定 서비스 業 産業實態 調査」에 의하면, 1991년도에 2,159億 5,500萬엔으로 처음으로 2,000億엔 대를 넘어섰다. 1,000億엔 대가 1985년이므로, 6년째에 접어들어 倍로 증가되었다.

3. 데이터베이스의 流通機構

데이터베이스 서비스는 몇개의 기업이 역할을 分擔하여 시행하고 있다. 데이터베이스 서비스의 업종에는 ①프로듀서, ②디스트리뷰터, ③인포메이션 브로커, ④에이전트, ⑤네트워크 業者 등이 있다.

프로듀서는 文字 그대로 데이터베이스를 構築(作成)하는 企業(機關)이다. 디스트리뷰터는 프로듀서가 작성한 데이터베이스를 通信回線에 실어서 利用者에게 제공하는 기업이다. 이 두 가지의 업종은 데이터베이스를 만들어 배포하는 기본적인 업무를 擔當하고 있어 가장 중요한 업종이다.

인포메이션 브로커는 代行檢索業으로 호칭되기도 한다. 디스트리뷰터와 계

약해 있지 않은 이용자 등을 위하여 데이터베이스에서 정보를 뽑아내어, 附加價値를 붙여서 제공하는 업종이다. 에이전트는 프로듀서나 디스트리뷰터를 위하여 販賣促進 업무를 대행하거나 利用者 支援 등을 하는 대리점을 말한다. 네트워크 업자는 디스트리뷰터에 통신 네트워크를 제공하거나, 게이트웨이에 의하여 디스트리뷰터끼리 연결시켜 이용자에게 다른 호스트의 데이터베이스를 제공하는 서비스이다. 이러한 업종은 프로듀서 및 디스트리뷰터의 일을 측면에서 支援하여 데이터베이스 서비스를 多樣化하는 점이 중요하다.

그 이외에도 간접적으로 데이터베이스 서비스에 관여하고 있는 企業群이 있다. 예를 들면, 호스트 컴퓨터나 端末機를 제조하고 있는 메이커가 있다. DBMS(database management system)나 通信 소프트웨어 등의 관련 소프트웨어를 작성하고 있는 소프트웨어 하우스도 있다.

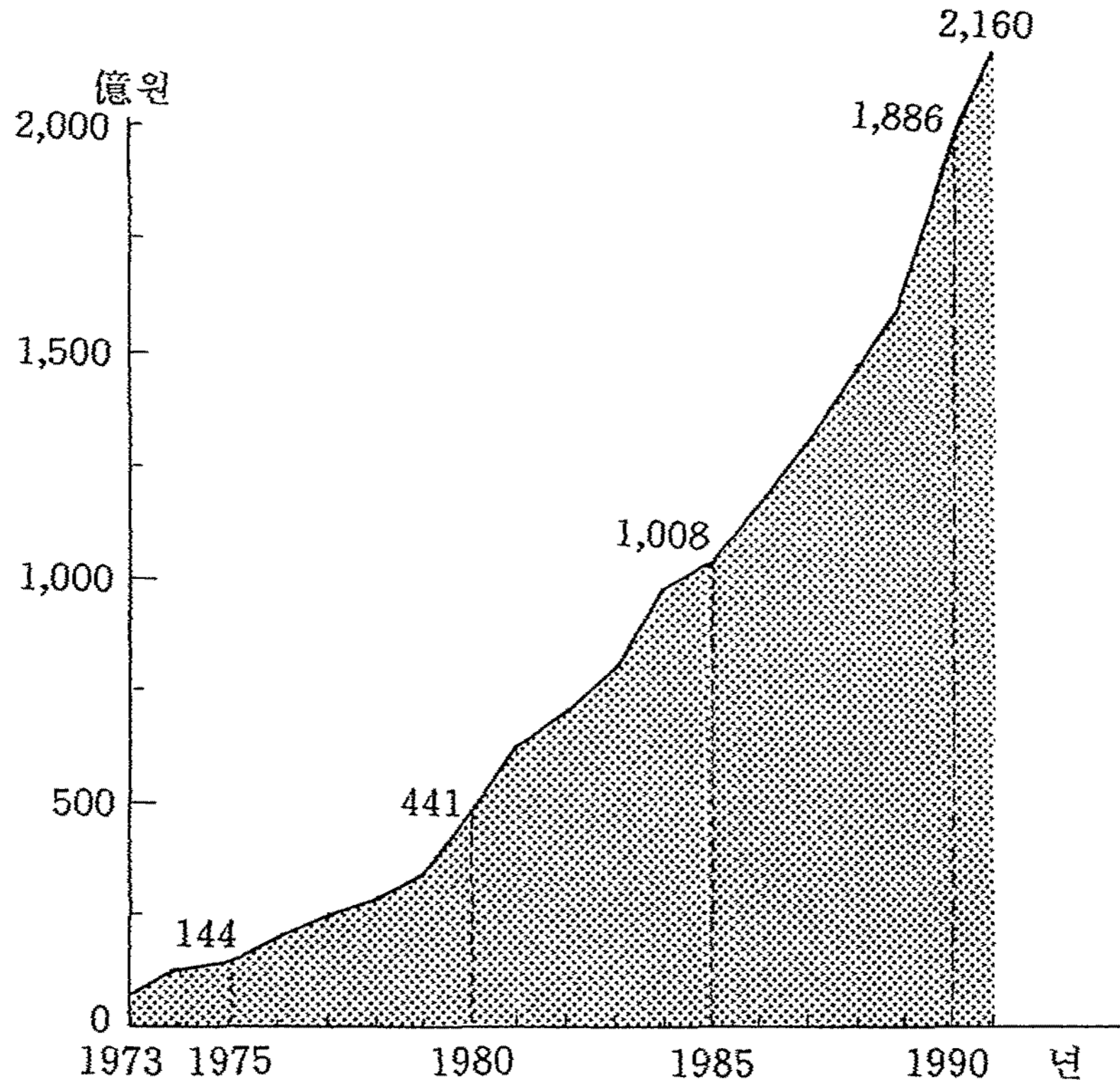
데이터베이스의 構築, 運用, 利用이라는 일련의 흐름은 이와 같이 다양한 企業群이 역할을 분담한 流通機構에 의하여 지탱되고 있다. 데이터베이스에서 앞서고 있는 美國에서는 프로듀서의 상당수가 이에 참여해 있고, 동시에 디스트리뷰터와의 分業體制가 명확히 구분되어 있다. 그러나 日本은 프로듀서와 디스트리뷰터의 업무가 같은 기업에서 兼業하고 있는 사례가 많다.

4. 데이터베이스의 流通狀況

日本에서 이용되는 데이터베이스의 種數는 「데이터베이스 臺帳總覽」에 의하면, 1991년도에 2,686(實數 베이스)種에 이르고 있다. 同 臺帳은 1983년에 通産省이 데이터베이스의 편람으로 발행한 것이다. 그 당시는 겨우 679種의 데이터베이스 뿐이었으나, 1985년도에는 1,000種을 넘었고, 1989년도에는 2,000種을 돌파하였다.

이 臺帳에는 日本의 데이터베이스 프로듀서나 디스트리뷰터가 構築 또는 提供하고 있는 데이터베이스가 수록되어 있다. 따라서 복수의 디스트리뷰터가 해외의 동일한 데이터베이스를 중복으로 登錄한 경우도 있다. 이 중복분을 제외한 것이 「實數」이고, 그대로 중복으로 계산한 것이 「收錄數」이다. 收錄數는 1991년도에 3,768種에 이르고 있다. 또한 日本에서 이용되는 데이터베이스의 거의가 해외에 있는 호스트에 저장된 외국계 데이터베이스이다. 현재로선 通信回線을 통하여 손쉽게 온라인이 이용됨으로 당연한 것이기도 하다. 그와 관

〈圖 5〉 데이터베이스 서비스 業의 연간 賣出額 推移



單位：億圓

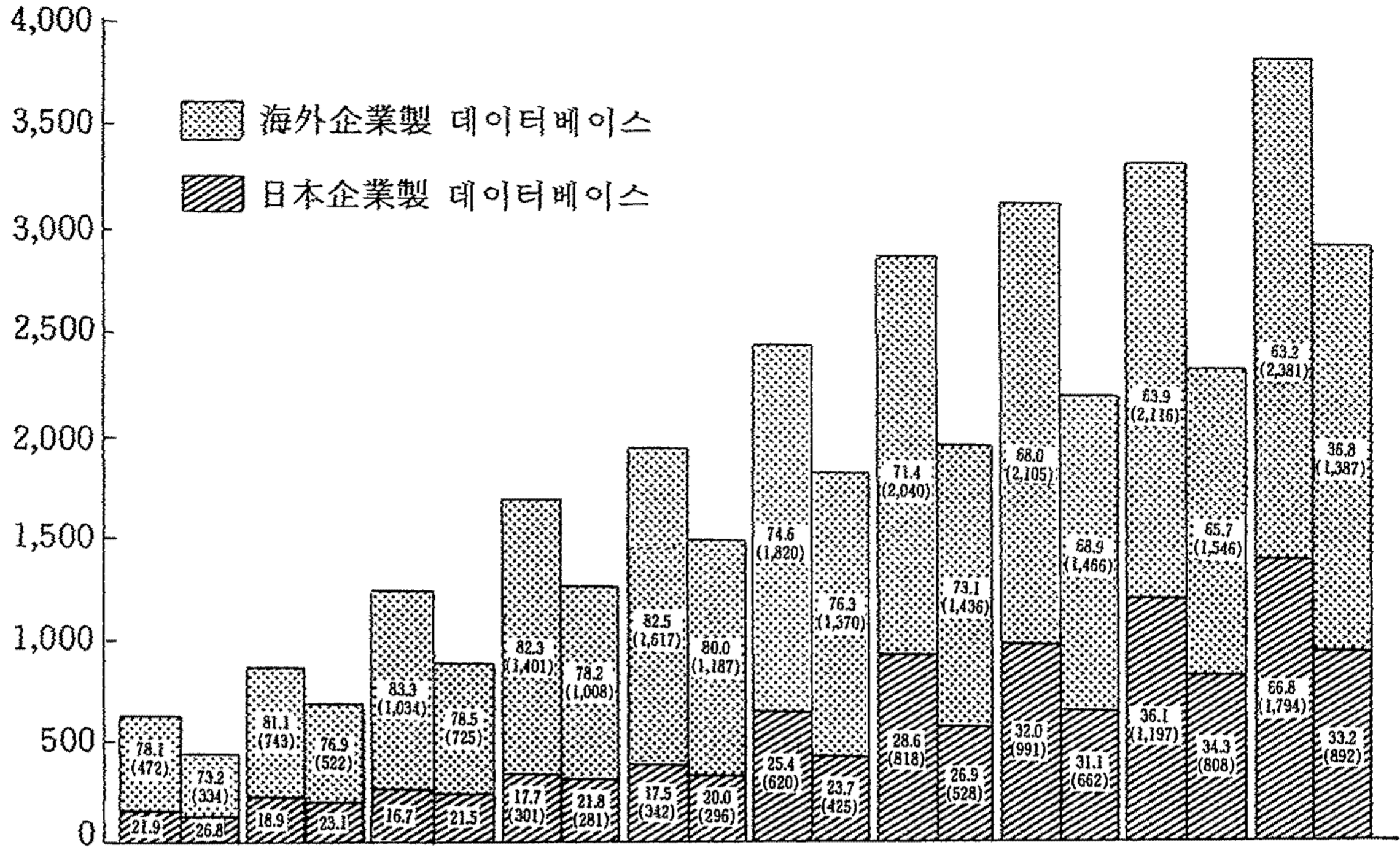
1973	1974	1975	1976	1977	1978	1979	1980	1981	1982
76	130	144	121	238	271	316	441	607	523
1983	1984	1985	1986	1987	1988	1989	1990	1991	
787	967	1,008	1,143	432	1,063	1,576	1,886	2,160	

資料：通産省, 「特定서비스 産業實態調査」에서 작성.

런하여 1991년도 實數의 66.8%는 외국제이다. 그러나 최근엔 國産 데이터베이스도 착실히 증대되었다. 1989년도에는 國産이 처음으로 전체의 30% 이상 (實數로 662種)이 되었으나, 1991년도에는 892種이 되었다(〈圖 6〉 참조).

日本의 데이터베이스에 관련된 成長推移를 보기 위하여 1983년도의 실적을 100으로 하여 1989년도와 비교하면, ①데이터베이스 賣出額은 2.7배, ②데이터베이스 實數는 3.9배, ③國産 데이터베이스 수는 5.7배, ④참여 기업수는 3.8배로 되어 있다(〈圖 7〉 참조).

〈圖 6〉 日本에서 이용되는 데이터베이스 實數와 收錄 데이터베이스 數의 推移

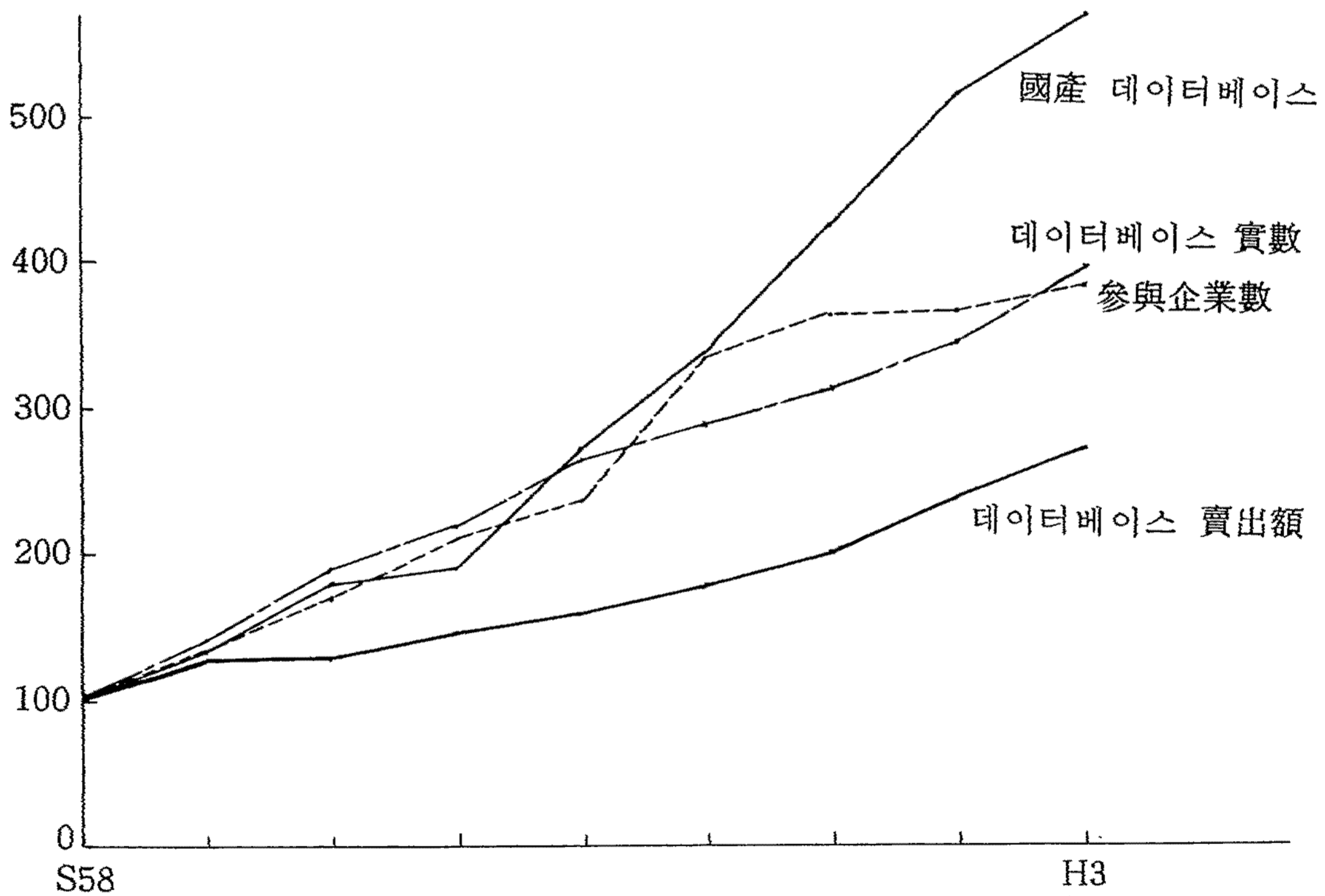


單位：億 원

年度	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
DB實數	456	679	924	1,289	1,483	1,795	1,964	2,128	2,354	2,686
收錄DB數	604	916	1,242	1,702	1,959	2,440	2,858	3,096	3,313	3,768

資料：通産省, 「데이터베이스 臺帳總覽」에서 作成.

〈圖 7〉 데이터베이스 관련의 成長推移(1983年을 100으로 한 경우)



資料：通産省, 「데이터베이스 臺帳總覽」; 「特定서비스 産業實態調査」와 근거하여 作成.

5. 분야별 流通狀況

日本에서 이용되는 데이터베이스의 분야별 分布를 보면, ①비즈니스 39.5%, ②自然科學·技術 29.6%, ③一般 26.8%, ④社會·人文科學 3.5%로 되어 있다(1991년도 實數 베이스). 이 분포 상황은 데이터베이스의 成熟度를 나타내는 파라미터이다.

데이터베이스는 그 발생의 역사를 보아도 당초엔 科學技術論文의 書誌的 事項이 주류이었다. 따라서 科學者나 學者, 엔지니어 등의 전문가가 주체였다. 그 후의 技術革新, 특히 記憶 미디어의 大容量化와 저렴化에 따라, 비즈니스 關聯記事나 新聞의 全文이 제공케 되었다. 이러한 경향에 연동하여 데이터베이스의 분야는 科學技術로부터 비즈니스로, 또한 이용자는 專門家로부터 일반 비즈니스로 多樣化 되어왔다(〈表 1〉 참조).

그와 관련해서 日本에 있어서도 일반 分野에 분류되어 있는 「新聞·雜誌·뉴스」와 「人物·機關情報」를 비즈니스 분야로 셈하면, 비즈니스는 1986년 이후 일관하여 50% 이상을 유지하고 있다. 보다 자세한 分野에서 보면, 流通 데이터베이스 數가 많은 분야의 톱 5는 아래와 같다(괄호 內는 1991년도의 實數).

〈表 1〉 데이터베이스 서비스 發展段階(美國)

段	階	局	面
1	1960년대	商用 以前	<ul style="list-style-type: none"> * 정부의 강력한 支援 * 助成 * 國家 프로젝트
2	1970년대	商用化(技術主導)	<ul style="list-style-type: none"> * 科學技術分野 주류 * 商用 서비스의 출현, 본격화 * 利用者는 기술자, 학자 등 專門家
3	1980년대	市場擴大(수요 主導)	<ul style="list-style-type: none"> * 비즈니스 분야 主流 * 競爭의 격화 * 利用者층은 비즈니스 머신으로 擴大
4	1990년대	市場擴大(퍼스널化)	<ul style="list-style-type: none"> * 취미, 레저, 生活情報의 확충 * 利用者는 個人으로 擴大

資料: 데이터베이스振興센터(DPC).

①「新聞·雜誌·뉴스」(281), ②「企業財務·企業情報(外國)」(238), ③「醫學·藥學·生命學·生物」(164), ④「金融·證券·換」(156), ⑤「市場·商品」(130).

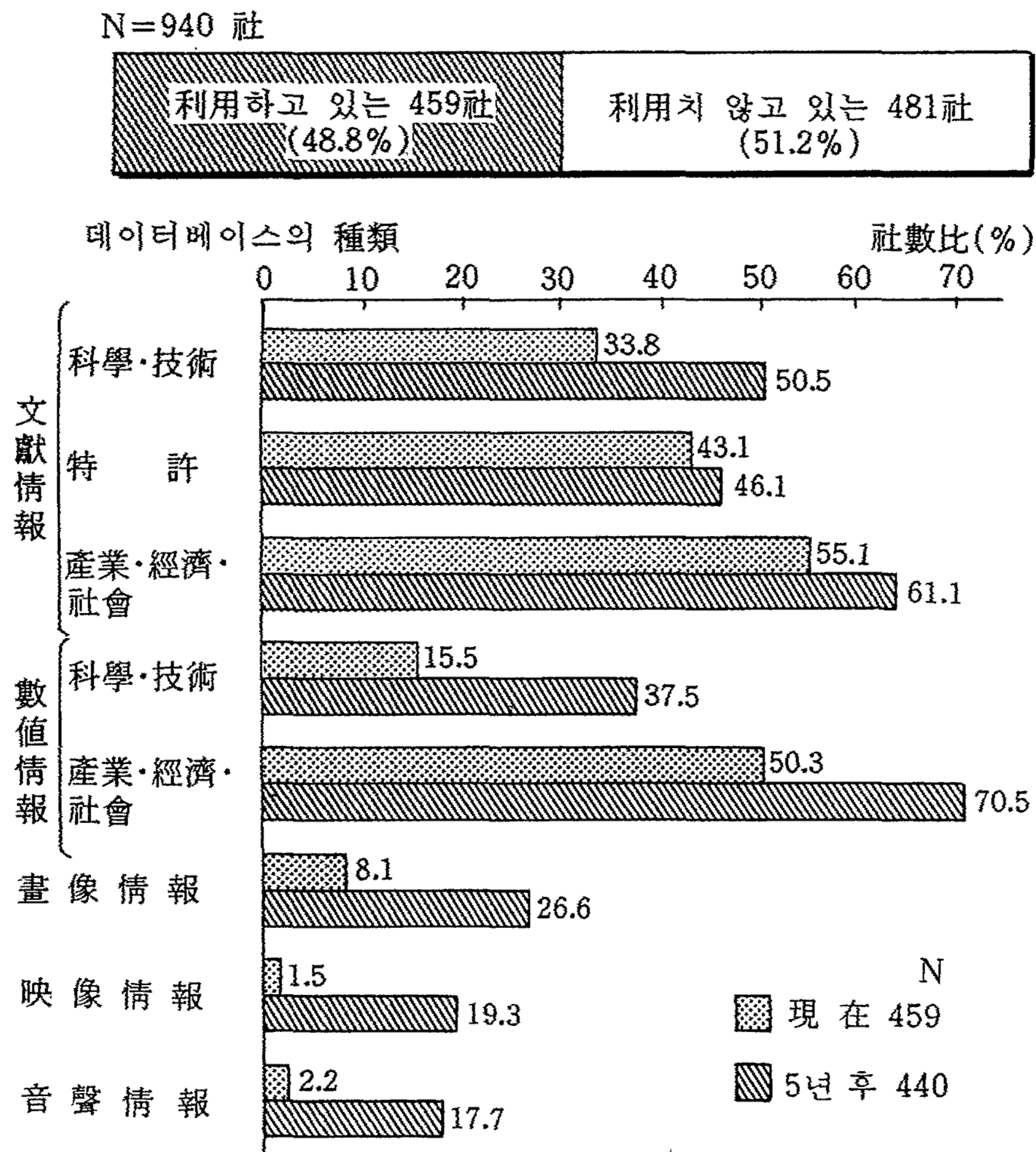
또한 기타의 분야에서 데이터베이스 수가 많은 것은 「人物·機關」, 「法律」, 「特許」, 「化學」, 「電氣·電子·情報」, 「經濟(매크로)」 등이다.¹⁾

IV. 데이터베이스의 利用狀況

1. JIPDEC 調査

(財)日本情報處理開發協會(JIPDEC)가 1991년에 실시한 調査에 의하면, 日

〈圖 8〉 商用 데이터베이스의 利用狀況(JIPDEC 調査 1991년)



資料:(財)日本情報處理開發協會, 「情報化白書」, 1992.

1) 本 講座는 앞으로 主要分野의 데이터베이스의 活用に 대하여 분야마다 시리즈로 다루어 소개한다.

본의 기업이 商用 데이터베이스를 이용한 것은 회답 940個社の 48.8%였다.

가장 많이 이용된 것은 文獻·數値를 불문하고, 産業·經濟情報였다. 이 경향은 5년 후에도 변하지 않을 것으로 예측된다. 단, 5년 후에는 畫像, 映像, 音聲 등의 데이터베이스 이용이 증대될 경향이 나타나 소위 멀티미디어 데이터베이스에의 기대가 높다. 현재의 CD-ROM 데이터베이스는 그 先驅者的 위치에 놓여 있다고도 말할 수 있다(〈圖 8〉 참조).

2. DPC 調査

(財)데이터베이스 振興 센터(DPC)는 매년 日本의 데이터베이스의 利用狀況을 파악하기 위하여, 「데이터베이스 서비스에 관한 利用者の 意識調査」를 실시하고 있다.

1991년의 同 調査에 의하면, 回答 774個社中 데이터베이스를 이용하고 있는 기업은 514個社로 66.4%였다. 또한 大企業은 70.5%가 이용하고 있음에 대하여 中小企業은 57.4%가 이용하고 있었다.

同 調査에서 日本의 데이터베이스 이용의 경향을 살펴보면, 아래와 같다.

(1) 利用金額

日本企業(公共 서비스 포함)의 데이터베이스의 年平均 利用額數는 3,152萬 엔(1990년도 실적)이다.

이 가운데 國產 데이터베이스가 73% 이상을 차지하고 있고 企業別로 보면, 大企業은 年平均 4,488萬엔, 中小企業은 331萬엔으로 격차가 크다.

(2) 利用形態

利用形態 중 ①온라인, ②온라인·오프라인 公用, ③오프라인의 回答比率은 각각 48.8%, 49.6%, 1.6%였다.

이를 보건대 오프라인(배치)만의 이용은 대단히 적고, 온라인 이용이 主流를 이루고 있음을 알 수 있다.

對前年比 變化는 온라인·오프라인 公用만이 증가되고 있다. 즉, CD-ROM 데이터베이스 등의 출현으로 앞으로 온라인·오프라인 형태가 共存하여, 제각기 補完的 역할을 다하게 되리라고 생각된다.

(3) 入手情報

商用 데이터베이스로부터 입수하고 있는 정보로서, 國產 데이터베이스에는 ①新聞·雜誌記事, ②企業情報, ③雜誌, 論文, 書誌情報가 우위를 점하고 있다.

해외 데이터베이스는 論文書誌, 特許, 新聞·雜誌記事의 순이다.

(4) 오프라인 利用의 形態

오프라인(배치) 이용에 있어 가장 많은 것은 디스트리뷰터로부터의 出力結果를 FAX 또는 郵便으로 받는 形態(回答企業의 58%)이다.

이어서 磁氣 테이프, 플로피 디스크, CD-ROM으로 구입하는 形態(36%)로 되어 있다.

특히, CD-ROM 데이터베이스의 購入條件은 1991년도에 前年對比 3.5 포인트 증대해 있다.

(5) CD-ROM의 이용

CD-ROM 데이터베이스에서 흔히 이용되고 있는 분야는 ①圖書·刊行物情報(CD-ROM 이용기업의 42%), ②醫學·藥學·바이오·化學(동 36%), ③辭典·디렉토리(동 22%), ④신문·잡지(18%)로 되어 있다.

CD-ROM 데이터베이스를 이용하고 있는 이유로는 ①空間 절약, ②通信料금이 불필요, ③일시불 固定料金, ④檢索 訓練에 좋다는 점 등이 있다.

V. 結 論

비즈니스에 있어서건 생활의 현장에 있어서건 정보는 不可避한 것이다. 그 때에 따라 필요한 정보를 이끌어 낼 수 있게끔 정보의 入手 루트를 다양하고도 충실하게 해두는 것이 중요하다.

특히, 商用 데이터베이스는 대가를 지불하면 누구든 자유롭게 이용된다. 바꾸어 말하면, 데이터베이스의 利用如何에 따라서 정보입수에 격차가 생기게 된다.

단, 데이터베이스는 「데이터」의 寶庫에 지나지 않는다. 데이터베이스에서 데이터를 끌어내어 이를 「情報」로 轉化하는 것은 어디까지나 利用者 자신이다.

그런 까닭에 檢索端末의 操作은 첫 걸음에 불과하다. 중요한 것은 이용자의 主題知識에 관한 깊이이다. 그 지식의 깊이에 따라서 데이터베이스로부터 「情報」를 끌어낸다고 하여도 지나친 말이 아니다.

더욱 중요한 것은 끌어낸 정보의 分析力과 活用力이다. 이를 갈고 닦아두지 않으면, 정보를 액션(判斷·行動)에 연결시킬 때 어긋나게 되어 버린다.

비즈니스 社會에 사는 우리들은 제각기의 部署에 있어 모두가 專門領域을 갖고 있다. 바꾸어 말하면, 데이터베이스의 이상적인 엔드 유저가 될 수 있는 위치에 있다. 이 강좌가 데이터베이스 활용의 길잡이가 되었으면 한다.

〈參考文獻〉

- (財)데이터베이스振興센터(DPC), 「데이터베이스 白書」, 1992.
- 通商産業省 産業構造審議會 情報産業部會, 「2000년의 情報産業 비전」, 1987.
- 通商産業省, 「平成3年度版 特定 서비스 産業實態調査」(平成 4년 12월).
- 通商産業省, 「데이터베이스 臺帳總覽」, 平成3年度版/平成 4년10월.
- (財)日本情報處理開發協會(JIPDEC), 「情報化白書」, 1992.
- 데이터베이스(DINA), 「데이터베이스入門 세미나 텍스트」.