

전세계 DB산업 현황

(The State of Databases Today : 1993) ③

차 례

| |
|--------------------------|
| I. 개요 |
| II. 데이터베이스 레코드 |
| III. 온라인 검색 |
| IV. 데이터베이스 (표현형태별 분석) |
| V. 지역 및 국가별 분석 |
| VI. 주제별 분석 |
| VII. 정보제공자별 분석 |
| VIII. 정보유통자별 분석 |
| IX. 데이터베이스세계의 미래상 |

“데이터베이스 세계는 역동적이다. 다시 말하면, 매년 성장하고 정제될 줄을 모른다는 것이다. DB의 유통과 접근을 위한 메체들의 발전잠재력 덕분에 DB 이용자들의 관심이 더욱 높아지고 있다.”

VI. 주제별 분석

대부분의 DB사용자들은 주제별 분류를 가장 중요시 하는데, 왜냐하면 이용하고자 하는 DB의 선택에 있어서 주제별 분류가 가장 중요하다고 생각하기 때문이다. 예를 들어, 검색대행자나 일반사용자가 화학정보를 찾고자하면, 그는 화학관련 DB를 선택할 것이다. 즉, DB의 형태나 제공매체는 두번째 문제라는 것이다. 검색 유형별로는 全文이나 통계DB가 필요하다고 할 것이다. 만약 선택된 주제항목에 아무것도 존재하지 않는다면, 주제항목에 있는 서지정보가 전문이나 전산용지로 출력된 통계정보를 대신할 것이다.

대부분의 DB들은 2개 혹은 그 이상의 일반적인 주제에 의해 분류될 수 있다(표5 참조). 예를 들어 BioBusiness DB는 건강/생명과학과 비즈니스 항목에 동시에 분류될 수 있다. 표5는 DB의 수와 비율보다는 DB의 주제별 분류에 중점을 두고 있다. 만약 1992년을 기준으로 비즈니스 분야로 분류된 DB의 비율을 알고싶다면, 주제별로 분류된 전체 DB수(7,871)로 비즈니스분야의 DB수(2,624)를 나누면 된다. 그러면 33%가 나오는데 이는 디렉토리에서 다루는 전체DB의 1/3이 비즈니스분야라는 것을

<표 5>

주제별 DB수 및 비율

| 주제별 분류 | 1988 | | 1989 | | 1990 | | 1991 | | 1992 | |
|----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 수 | (%) | 수 | (%) | 수 | (%) | 수 | (%) | 수 | (%) |
| 비즈니스 | 1,815 | (34) | 1,687 | (33) | 1,956 | (33) | 2,101 | (33) | 2,624 | (33) |
| 소비자 관련 | 186 | (3) | | | | | | | | |
| 일 반 | 301 | (6) | 327 | (6) | 416 | (7) | 450 | (7) | 700 | (9) |
| 건강/생명과학 | 433 | (8) | 576 | (11) | 651 | (11) | 690 | (11) | 728 | (9) |
| 인류(Humanities) | 84 | (2) | 184 | (3) | 216 | (4) | 248 | (4) | 314 | (4) |
| 법 | 441 | (8) | 447 | (9) | 531 | (9) | 574 | (9) | 885 | (11) |
| 전문학문 | 29 | (<1) | 335 | (7) | 368 | (6) | 366 | (6) | 296 | (4) |
| 뉴 스 | 428 | (8) | 186 | (4) | 233 | (4) | 291 | (4) | 385 | (5) |
| 사회과학 | 460 | (9) | 393 | (8) | 418 | (7) | 253 | (7) | 447 | (6) |
| 공학분야 | 1,184 | (22) | 996 | (19) | 1,154 | (19) | 1,210 | (19) | 1,492 | (19) |
| 총 계 | 5,361 | (100) | 5,131 | (100) | 5,943 | (100) | 6,383 | (100) | 7,841 | (100) |

* 주 : 한 DB가 여러 주제분류에 포함될 수도 있다.

나타낸다. 어떤 분야가 전체 가운데 아주 적은 부분을 차지한다면, 그 주제항목은 따로 분류되지 않았다.

비즈니스관련 DB들이 가장 많고 다음은 과학/기술/공학분야이며, 세번째로 많은 분야는 법관련 DB들로 1989년이래 거의 2배로 증가했다. 그리고 4번째는 건강/생명과학분야로 법분야와 비슷한 수를 보이고 있다. 상당수의 비즈니스DB들이 법, 뉴스 관련 DB들과 중복분류되었다. 소비자관련 항목들은 소비자들의 관심을 유도하고 그들이 이용할 수 있는 정보들을 분류하였다. 이것은 소비자들과 관련된("about") 정보와는 구별된다. 소비자와 "관련된" 대부

분의 DB들은 시장연구 혹은 인구통계와 관련된 내용이어서 비즈니스 혹은 사회과학분야로 분류되었다. 소비자관련 분류부분은 1989년에 중단되었다.

DB를 살펴보는 또하나의 방법은 유통과 접속을 위해 사용된 기록매체의 종류에 따른 분류방법이다.<표5>는 다양한 매체로 기록된 DB의 수와 그 비율을 나타낸다. 1992년의 경우 6,998개의 Entry DB를 기록매체별로 분류해 결합시켰더니 8,495가지가 되었다. 전체 DB중에서 온라인 형태는 65%, CD-ROM 형태는 15%, 디스켓 8%, 마그네틱 테이프 7%, 배치형태 5%, 그리고 기타 1%이하였다.

<표 6>

DB 유통과 이용에 이용되는 매체별 분석

| 매 체 | 1989 매체-DB 갯수 | | 1990 매체-DB 갯수 | | 1991 매체-DB 갯수 | | 1992 매체-DB 갯수 | |
|---------|------------------|-------|------------------|-------|------------------|-------|------------------|-------|
| | 수 | DB수 | 수 | DB수 | 수 | DB수 | 수 | DB수 |
| 온라인 | 3,524 | 2,711 | 4,018 | 3,041 | 4,170 | 3,200 | 5,486 | 4,519 |
| 배치형태 | 999 | 769 | 1,252 | 948 | 1,321 | 1,014 | 389 | 320 |
| CD-ROM | 433 | 333 | 715 | 541 | 1,019 | 782 | 1,321 | 1,088 |
| 디스켓 | 478 | 358 | 626 | 474 | 695 | 533 | 676 | 557 |
| 마그네틱 테잎 | 787 | 605 | 906 | 686 | 954 | 732 | 584 | 481 |
| 기타 | - | - | - | - | - | - | 39 | 33 |
| 총 계 | 6,221 | 4,786 | 7,517 | 5,689 | 8,159 | 6,261 | 8,495 | 6,998 |

* 주 : 한 DB를 몇개의 매체에서도 이용가능함.

VII. 정보제공자 (Producers) 별 분석

정보제공자는 때로는 DB유통자(Publisher)라고 불리기도 한다. 왜냐하면 그들이 공공적으로 유통("make public")시키기 때문이다. 어떤이들은 전자적인 형태로만 데이터를 유통시키는 반면, 자신들의 DB를 인쇄물 형태로 출력한 것을 유통시키는 곳도 있어 이들을 전통적인 의미에서 유통자(publishers)라고 말한다. DB제공자들은 DB의 내용을 결정하여 이를 만들어내는 직접적인 당사자들이다. 대부분의 정보제공자들은 임대 혹은 계약에 의해 개별기업이나 유통자(vendors)에게 자신들의 DB를 제공한다. 유통자는 이용료를 받고 이용자들에게 검색서비스를 제공한다. Mead Data Central, U.S. National Library of Medicine,

그리고 DRI/McGraw-Hill같이 정보제공자이면서 유통자를 겸한 곳이 점차 증가하고 있는데, 이들은 자신의 DB뿐만 아니라 다른 제공자의 DB로부터 배치 혹은 온라인 형태의 검색서비스를 제공하고 있다.

정보제공자는 정부관련기관, 기업/산업체, 비영리단체(학원포함), 그리고 이들의 혼합형태 등 하부구조안에서의 위치에 따라 분류할 수 있다. 1960~70년대에는 대부분의 DB가 정부관련기관, 예를 들어 미항공우주국(NASA), 미국 원자력위원회(현재의 에너지부)에서 제작되었고, 1960~70년대 후반에는 NFP DB가 주로 전문가집단을 중심으로 그 중요성에 따른 이용이 늘어나면서 두각을 나타냈다. 과학분야에서 정부관련DB와 NFP DB가 중요한 자원으로 여겨졌지만 이들의 DB수는 약간

<표 7>

정보제공자별 DB 점유율

| 정 보 제 공 자 | 1977 % | 1985 % | 1988 % | 1989 % | 1990 % | 1991 % | 1992 % |
|-------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 정부/관련기관 | 56 | 21 | 20 | 17 | 17 | 16 | 15 |
| 기업/산업체 | 22 | 57 | 65 | 68 | 68 | 70 | 75 |
| 비영리단체(학원포함) | 22 | 11 | 13 | 12 | 12 | 12 | 9 |
| 기 타 | 0 | 11 | 2 | 3 | 3 | 2 | 1 |
| 총 계 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

늘어도 시장점유율은 하락하고 있다. 정부 관련 DB는 1977년과 1992년 사이에 56%에서 15%로 줄어들었고, NFP DB도 같은 기간 동안 22%에서 9%로 감소되었다. 반면 상업적 DB들은 수적으로 계속 늘어 1977년의 22%에서 1992년에 72%로 증가해왔다 (표7참조).

어떤 DB들은 복합적 성격의 기관에서 제공되거나 여러기관에서 정보를 제공해서 만들어진 것도 있다. 복합적 성격의 대표적인 것이 다수 국가적 성격의 기관에서 제공하는 것인데, 예를 들면 1개이상의 국가를 대변하는 기관으로 UN, 유럽경제공동체(EEC)가 있다. 정보제공자의 성격은 시간에 따라 많은 변화를 맞았는데 이에 관한 가장 최초의 통계는 1977년에 만들어졌다. 1985년부터 1991년 사이에 점유율은 별로 변화가 없었고 기업/산업체 분야에서 1992년에는 전년에 비해 5%가 늘어났다 (표7참조). 정부관련 DB들이 점유율에서는 줄어들었지만 많은 상업적 DB들이 정부의 자료를 바탕으로 제작되었다는 점을 유의

할 필요가 있다. 예를 들어 인구통계조사는 엄청난 비용을 들여 정부에서 모은 자료인데, 상업적 DB구축자들은 이 자료를 받아 부가가치를 높여 상업적 DB로 가공, 재판매할 수 있었다는 것이다.

현재 유통되는 전체 DB수가 증가하는 것과 마찬가지로 개별 정보제공자 그룹별 DB수도 증가해왔다. 위에서 보듯이, 상업적DB entry의 수도 1977년의 82개에서 1992년에는 5,224개로 급격하게 증가했다. 학원을 포함한 비영리단체의 DB entry도 1977년에 82개에서 1992년에는 633개로 증가했고 같은 기간 정부관련 DB의 수도 1977년의 204개에서 1992년에는 1,039개로 늘어났다.

VIII. 유통자 (Vendors) 별 분석

유통자는 DB에 가치를 부여하여 유통시키고 온라인이나 배치형태의 검색서비스를 제공하거나 이용자에게 CD-ROM 타이틀을 배포하는 업체들을 말한다. 수많은 유



통자들은 DB유통외에도 여러가지 서비스를 제공하기 때문에 유통자(Vendors)라고 불리는 것을 싫어한다. 왜냐하면 그들은 DB에 가치를 부여하고 이에 접속하는 기술을 제공하기도 하기 때문이다. DB에 가치를 부여하는 방법은 몇가지가 있다. 즉, DB개발>Loading), 검색 S/W의 기능 향상, 그리고 온라인 문서주문과 같은 연관서비스, 정보선택제공(SDI)등이 그것이다. DB 유통자에 대해서는 과거부터 지금까지 전세계 여러지역에서 다양하게 언급되어 왔다. '1960년대 미국에서는 테일 판매상들이 유통자로 간주되었고, 유럽에서는 유통자들이 "Host"라고 불리었다. "유통자"라는 단어에 대한 거부감이 있었음에도 불구하고, DB를 검색하는 고객이나 이용자들에게는 선호되는 용어가 되었다. 컴퓨터를 이용하여 검색 가능한 DB의 검색서비스를 제공하는 업체를 대상으로한 논문이나 조사보고서에서는 가장 흔히 사용되는 용어가 되었다.

DB를 선별해서 분류하는 주요한 유통자에는 DIALOG Information Services, BRS Information Technologies, ORBIT Search Service, U.S. National Library of Medicine,

Questel, PFDS Online(과거 Pergamond Financial Data Services), Data-Star가 있다.全文DB 제공업체들로는 Information Services, 그리고 DataTimes Corporation이 있다. 계량경제DB를 제공하는 유통자에는 The WFA Group(과거 Wharton Electronic Forecasting Associates and Chase Econometrics), Reuters Information Services(캐나다) Ltd. (과거 I.P. Sharp Associates), 그리고 DRI/McGrawHill이 있다. 또한 계량적인 과학관련 DB를 유통하는 업체는 CAN/SND, STN International이 있다. 소비자 중심적인 DB를 유통하는 업체에는 Compu-Serve Information Service, PRODIGY가 있다.

이글에서 분석한 유통자의 수는 1,438개였다.

IX. 데이터베이스세계의 미래상

데이터베이스세계는 역동적이다. 다시말하면, 계속 변화하며 매년 성장하고 아직도 정체될 줄을 모른다는 것이다. DB의 유통과 접근을 위한 새로운 매체들의 잠재력 덕분에 DB의 이용자들의 관심이 더욱 높아가고 있다. 특히, CD-ROM DB의 개발과 유통으로 말미암아 그동안 비용문제가 접근의 장애물로 여겨졌던 대학에서 컴퓨터를 이용한 DB이용이 크게 늘어나고 있다. 그와 동시에, CD-ROM으로 유통된 DB들에 대한 온라인 검색이용이 줄어들고 있다. 그러한 감소추세는 ERIC, PsycINFO, NTIS, 그리고 대학에서 많이 이용되었던 MEDLINE에서 특히 두드러지게 나타나고 있다. [] <번역·정리 - 이석기>