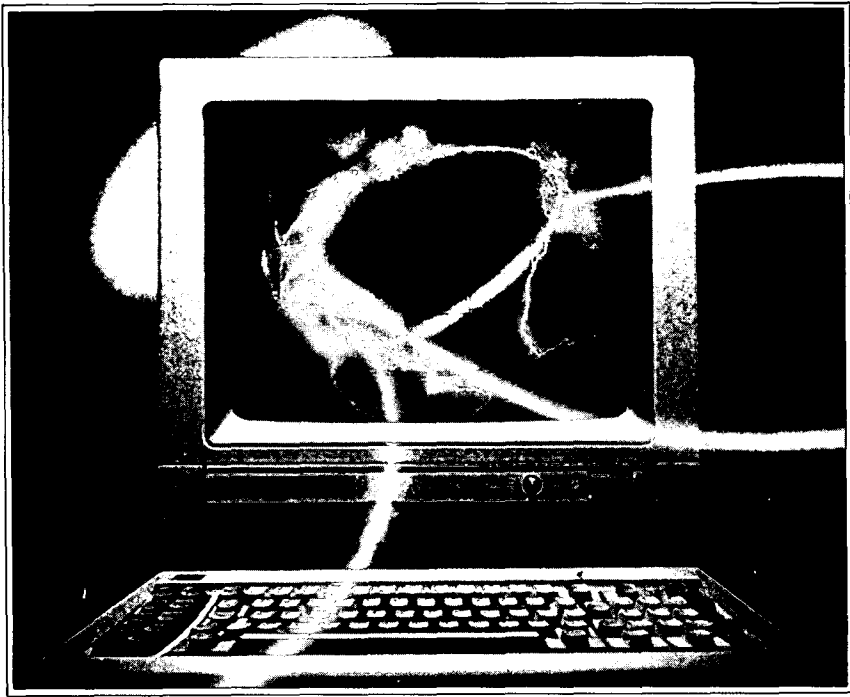


해외DB분석을 통해본 국내DB개발방향



우리나라의 데이터베이스 산업이 아직 국가의 제반산업에 비해서 상당히 뒤져있지만 아울러 향후 수 년내에 정보선진국으로의 도약이 국가적 과제가 되어 있는 현실에서 국내DB의 올바른 개발방향 정립은 중요한 일이 되고 있다.

장 재 경

1. 개 요

이 미 세계적으로 촌음을 다투어 가며 정보가 쏟아지고 있으며 이러한 온라인정보는 언론매체를 통해서 접하기 전에 선물시장, 주식시장 등에 영향을 미치는 한편, 세계 구석구석의

정보도 지구촌의 거리를 느낄 수 없을 만큼 쉽게 접할 수 있게 되었고, 간단한 전화번호 정보에서부터 개인이나 기업의 이해관계에 관련한 정보까지 필요에 따라 유용하게 활용할 수 있게 되었다.

데이터베이스는 이제 정보화 사회에서 국력의

단면을 상징하는 국가자원으로 인식되고 있다. 이는 기존의 천연자원이나 인력자원과 같은 국가 자원과는 달리 사용함으로써 가치가 고갈되지 않으며, 새로운 정보의 창출을 위한 시간, 노력을 놀랄만큼 단축시켜 주어 국가경쟁력을 향상시키는 몫을 담당하게 된다.

DB개발에 앞서 데이터베이스가 중요시 되고 있는 이유는 정보를 통하여 다른 이익을 창출해 내는 생산성에 있다고 할 수 있다. 다시 말해서 국가, 사회 기반 산업으로서의 데이터베이스 산업에 대한 기대는 국민에게 필요하고 유용한 정보를 신속하게 확보케 함으로서 국가 국제경쟁력을 제고한다는 것이다.

따라서 데이터베이스개발의 목적은 필요로하는 부가가치가 높은 유용한 정보를 발굴해 내는데에 있다고 하겠으며 그 개발기준을 다음과 같이 생각해 보았다.

- 정보이용자의 생산가치(이용가치)가 높은 분야의 DB개발
- 각 분야별 기존지식(신문 또는 책으로 배포된 지식)을 효과적으로 총괄한 DB개발
- 공간적 시간적 요인으로 분산되어 있는 현황정보를 통합한 DB개발
- 최신성과 신뢰성을 유지할 수 있는 DB개발

2 해외DB의 최신 동향

1960년대 소련의 소푸트닉 쇼크에 따라 미국정부의 전폭적인 지원으로 육성되기 시작한 MEDLINE, ERIC, AGRIS등과 같은 공공부분의 DB개발에 이어 세계 최초의 상용 DB인 DIALOG가 출현한 이래 미국은 DB산업의 선진국이 되어 전 세계의 DB시장을 석권하고 있다. 또한 정보선진국이라 할 수 있는 프랑스는 미니텔을 중심으로 국민생활에 기여하고 있는바 이는 사회주의적 경제, 사회구조속에서 효과적으로 성장할 수 있었으며, 일본의 경우 동양적인 사고방식으로 상대적으로 DB산업부분의 발전이 떨어졌으나

'80년대 후반에 이르러 본격적인 육성책인 DB준비금제도, 일본개발은행에 의한 출자제도 등의 정책에 힘입어 '90년대에 진입하면서 <표1>에서 일본의 정보화 지수를 보면 괄목할 만한 급격한 성장을 하게 되었다.

	한국	미국	일본	독일	영국	프랑스
'84	23	426	173	174	155	197
'85	29	475	212	227	190	236
'86	34	524	331	302	241	311
'87	46	582	428	380	301	372
'88	60	646	581	457	362	416
'89	81	723	748	546	444	478
'90	100	808	1028	695	542	579

표1 주요국의 정보화 지수

전세계 DB시장을 주도하고 있는 미국의 DB시장도 최근 국제화 경향과 발 맞추어 전체 이용량이 늘어나는 흑자로 기록하고 있어 정보화사회의 새로운 장이 서서히 열려가고 있음을 실감할 수 있다.

정보시장의 규모가 커짐에 따라 해외의 DB업체는 단순한 1, 2차 정보의 가공제작뿐만 아니라 보다 부가가치가 높은 3차정보 즉 예측정보, 분석정보 등에 대한 관심과 투자가 늘어가고 있다. 한편 컴퓨터 하드웨어 및 관련장비의 빠른 발전과 맞추어 새로운 DB제작 매체에 의한 신상품 개발도 볼 수 있다.

21 全文情報의 증가

최근 전자출판이라고 일컬어지는 CTS(Computerized Type Setting)출판 및 DTP(Desk Top Publishing)출판의 보급으로 정보의 전문을 자연발생적으로 DB화 할 수 있게 되었다. 해외의 대개의 신문을 포함해서 많은 저널이 온라인화하거나 전문(fulltext)을 제공하고 있는 경우가 늘어가고 있다. <표2>와 같이 약 2,000여 DB가 전문을 제공하고 있으며 이는 전체 DB의 약 3분

지 1수준이지만 최근 전문을 제공하는 DB의 수가 가속적으로 늘어가고 있는 추세이다.

이처럼 DB에서 전문을 제공하게 됨으로써 기존의 2차정보 DB에서의 경우 원문을 입수하는데 소요되는 시간 (약 2-3주 소요) 및 불편함이 크게 해소되어 정보이용의 편의를 배가하고 있으며, 또한 전문내에 있는 세부 내용까지 검색을 할 수 있기 때문에 기존의 DB에 비해서 전문에서 정교한 사실 또는 수치확인 까지 손쉽게 가능하게 되었다.

이로서 정보수집에 따르는 국가·지역간 장소의 제약 및 입수절차 및 지연시간에 따른 번거로움을 크게 극복하고 있으며 정보화 사회에서의 정보의 위력을 새삼 강조할 수 있게 되었다.

온라인서비스	전문정보제공 DB수
DIALOG	500
DATA-STAR	423
NEWS NET	366
MEAD DATA CENTRAL	246
WESTLAW	166
VU/TEXT	92
Data Times	82
Dow Jones	59
BRS	19
G-Search	10
Questel	9
STN	7

표2 온라인서비스에서의 전문정보제공DB수(1993.3 조사)

2-2 CD-ROM을 통한 原文情報の 증가

몇년전부터 세계DB전시회를 주도하고 있는 CD-ROM정보는 획기적인 기억용량과 빠른 액세스능력을 지니고 있어서 새로운 정보저장 매체이다. 아울러 CD-ROM은 통신망을 통해서 제공하기 어려운 그림, 도표, 사진 등을 포함한 원문을 그대로 수록할 수 있기 때문에 CD-ROM을

통한 원문정보를 제공하는 경향이 크게 증가하고 있다.

이는 최근 CD-ROM정보가 크게 주목을 받고 있는 이유 중의 하나이다. 즉 현미경을 이용한 조직생체사진을 찍 수 없는 의약, 생화학 정보의 경우, 또는 복잡한 화학 구조식을 그림으로 이해시키는 것이 효과적인 경우 또는 기술정보의 핵심이라고 할 수 있는 제법을 기술한 특허도면의 경우에는 그림, 사진은 매우 중요한 의미를 갖게 된다.

따라서 최근 의약학 분야에서 학술지를 발행하는 IP(출판사계열)들 간에 정보를 통합하여 User의 요구에 들어 맞는 정보만을 취사선택하여 이를 CD-ROM에 축적하여 공급하고 있는 새로운 대형 상품들이 나오고 있고, 각 IP들이 소유한 원문정보와 그 분야와 관련한 색인, 초록을 제공하는 권위있는 2차정보 데이터베이스를 통합하여 새로운 CD-ROM상품을 개발하고 있다.

아울러 DB에 문자정보 뿐만 아니라 음성까지도 함께 수록한 멀티미디어형 DB를 수록한 CD-ROM상품도 최근 선보이고 있으며 관련한 하드웨어로 빠르게 대중화되고 있다.

2-3 온라인 서비스업체의 통합, 대형화로 DB액세스가 편리

실제 해외 DB를 이용함에 있어서 각 온라인 서비스에 따라 검색 방법 및 제공되는 DB종류도 상이하기 때문에 검색에 있어서 제한이 있다. 그러나 최근 온라인 서비스가 통합·대형화 되는 경향이기에 때문에 User입장에서는 동일 온라인 서비스를 통해서 보다 다양한 세계 각국의 정보를 액세스할 수 있게 되었다.

- 미국의 DIALOG와 유럽의 Datastar의 통합
- EASYNET 서비스(11개의 온라인 서비스를 통합하여 1,000여 데이터베이스와 일시에 접속이 가능)
- OCLC의 방대한 정보원 및 상용DB를 통합한 EPIC, First Searching 서비스

표3 EASYNET이 제공하는 데이터뱅크현황

데이터뱅크명	국 명	정 보 내 용
DIALOG	미 국	비즈니스, 과학, 기술분야, 사회과학, 인문과학 등
BRS	미 국	국제뉴스, 방송, 신문기사정보, 의학, 교육, 엔지니어링 등
ORBIT	미 국	종합정보 제공 특히 과학기술 분야의 정보가 특히 강함
NewsNet	미 국	우주, 컴퓨터, 전자공학, 증권투자, 출판, 미정부법령 등
Pergamon	미 국	종합정보 제공 특히 비즈니스정보가 강함
Questel	프랑스	종합정보 제공 특히 유럽의 과학정보가 강함
G.Cam - Sinong	프랑스	프랑스 비즈니스, 정부관련 뉴스
DataStar	스위스	유럽 및 세계의 과학, 기술, 의학, 사회과학, 인문과학 등
Profile	영 국	세계 각국의 통신사, 라디오, 신문 등의 기사전문 정보
VU/TEXT	미 국	미국 지방신문, 저널 등의 기사정보
Wilson	미 국	교육, 비즈니스, 응용과학, 기술, 미술, 사회과학 등

24 IP의 전문화·대형화로 DB의 고급화

대개 공공DB의 육성을 선두로 시작한 각국의 DB산업에서 결국 상용DB의 종류가 늘어나서 DB산업의 주축을 이끌고 나가게 되는 것은 각종 상용DB들이 정보이용자의 요구를 면밀하게 파악하여 쉽게 대응하기 때문이라고 볼 수 있다.

미국의 경우 최근 많은 IP 또는 디스트리뷰터들이 세계적인 출판업체, 출판유통업체 혹은 글자의 언론매체 등의 계열회사로 바뀌어 가고 있는 사실이다. IP들이 대형화됨에 따라 투자의 규모가 커지고 DB도 고급화되어 가고 있다.

- 출판업체열의 IP
Pergamon사 (Maxwell Online Inc.)
McGraw-Hill사
Dun & Bradstreet사
Standard & Poor사
Predicast사
Marusen사 등
- 언론기관 계열의 IP
Knight-Ridder사(DIALOG Information Service Inc.)
Dow Jones 사 등

- 보험·금융 계열의 IP
H. & R. Block사(CompuServe Inc.)
Citibank

3. 국내DB 개발방향

우리나라 DB산업은 초기 단계로서 규모는 작지만 '91년을 기준으로 '92년에는 DB수는 58.5%, 이용자수는 178.9%, 매출액은 26.2% 증가하는 놀라운 성장상을 보이고 있는 고무적인 상태라고 할 수 있다.

	'90년	'91년	'92년	성장률(%)
DB수(개)	183	384	460	58.5
DB유통업체(개)	-	11	28	154.5
DB제작자수(명)	-	213	254	19.2
이용자수(천명)	45	290	350	178.9
매출액(억원)	470	567	748	26.2

표4 국내DB산업 추이

하지만 이러한 수량적인 증가에 비해 아직도 DB를 제작하는 데는 많은 어려움을 안고 있는 현실이다. 이를테면 아직도 정보에 대한 사회적 인식이 낮아서 시장이 협소하다는 점과 정보·자료 공개에 대한 의식이 미흡하다는 점을 들 수

있다.

그러한 여건에도 불구하고 보다 나은 DB개발을 위한 방향을 제시해 보면 다음과 같다.

3.1 금융, 경제, 신용, 여행, 예약 부문 등 專門상용DB 개발필요

일본의 데이터베이스진흥센터가 조사한 <표 5>에서 본다면 북미의 온라인 DB시장 매출액의 대부분을 금융·경제·신용정보가 41.6%를 차지함을 알 수 있다. 또한 여행 및 예약관련 정보에 관련한 매출도 17.9%에 이르고 있다. 한편 높은 성장성이 기대되는 분야는 제품정보, 상거래 서비

스 및 과학기술, 의료정보로 분석이 되고 있다.

또한 해외 상용DB를 분야별로 분포범위를 살펴보면 다음 <표6>과 같다.

분 야	시 장 규 모 (%)
금융 경제	24.6
여행·예약	17.9
신용 정보	17.0
마케팅·미디어	13.9
법률, 법규, 특허	5.1
과학기술·의료	4.9

표5 북미의 온라인 데이터베이스 시장 1990년

일 본			해 외		
비즈니스 360(54.5%)	경제	112(16.9%)	비즈니스 521(35.5%)	기업·재무	140(9.5%)
	기업·재무	75(11.3%)		경제	72(4.9%)
	금융·증권	70(10.6%)		금융·증권	70(4.8%)
	시장·상품	38(5.8%)		시장·상품	64(4.4%)
	전반	18(2.7%)		전반	44(3%)
	에너지산업	2(0.3%)		에너지·산업	36(2.5%)
	기타	45(6.8%)		기타	95(6.5%)
자연과학·기술 97(14.7%)	의학·약학	27(4.1%)	자연과학·기술 534(36.4%)	의학·약학	102(6.7%)
	특허	6(2.4%)		화학	87(5.9%)
	전반	15(2.2%)		전기·전자	77(5.3%)
	화학	7(1.1%)		전반	45(3.1%)
	전기·전자	4(0.6%)		특허	49(3.3%)
	기타	97(14.7%)		환경·공해	45(3.1%)
일반 185(27.9%)	신문·잡지	61(9.2%)	일반 344(23.5%)	신문·잡지	131(8.9%)
	인물·기관	30(4.5%)		전반	37(2.5%)
	전반	25(3.8%)		인물·기관	28(1.9%)
	건강·스포츠	16(2.4%)		행정·법률	95(6.5%)
	행정·법률	18(2.7%)		건강·스포츠	26(1.8%)
	사회·인문	16(2.4%)		기타	27(1.9%)
	기타	35(5.3%)		교육학	17(0.7%)
				인구·통계	4(0.3%)
		인문·사회	65(4.4%)		
		기타	2(0.5%)		

표6. 해외 분야별 상용DB분포(1989일본DB총량)

여기에서 보면 일본의 경우는 비즈니스 분야가 압도적으로 많고 과학기술 분야에서는 서구DB의 비중이 단연 높음을 알 수 있다.

한편 현재 국내DB산업의 경우는 규모면에서 작기도 하지만 매출액이 높은 비즈니스 등 전문 분야의 DB는 거의 전무한 실정이다. 해외의 많은 IP들이 기존에 정보를 취급하던 출판업에서 자연 발생적으로 성장을 볼 수 있는 바(예를 들면 기업체 관련 신용정보를 분석하여 출판보급을 하던 미국의 Dun & Bradstreet사 등은 세계 굴지의 기업관련 DB생산 IP가 되었고, 세계 각국의 특허 정보를 분석하여 책자형으로 보급하던 영국의 Derwent사도 현재 유명한 특허DB인 <WPI>로서 많은 이윤을 창출하고 있다.) 우리의 경우도 자연스럽게 전문분야의 정보를 DB로서 개발할 수 있는 가능성이 있다고 하겠다.

아직 우리에게 많은 전문분야의 정보가 책자형의 자료로 보급되었으나 이것이 DB화 되어 있지 못한 부분이 상당히 많이 있기 때문에 이것을 상용DB로 육성하는 것도 어렵지 않은 일이라 하겠다.

DB개발이 기존지식의 재가공이라는 측면에서 볼때 가공할 수 있는 국내에 축적된 기존지식의 분포를 파악과 동시에 국제화, 정보화 시대에 발생할 국내의 정보요구에 맞추어 국내 신규DB개발의 방향이 설 수 있을 것이다.

3-2 해외정보와 견줄 수 있는 고품질 정보의 개발

현재 서비스 되고 있는 해외DB는 TV뉴스, 드라마대본에서 부터 각종 이벤트 정보, 기업신용 분석정보, 투자상담 정보, 특정기업에 대한 추적 정보, 첨단과학기술 정보등 DB의 수는 약 7,000여개에 이르고 있다.

예를 들면 해외의 여행정보시스템의 경우 여행 정보 제공은 물론 항공권 예약(가장 저렴한 비행기 좌석을 찾아 예약해 주는 "Low Fare Finder"기능도 제공)도 해주며 호텔의 예약에 있

어 도심을 중심으로 한 반경 Km단위로 리스트를 제공하는 한편, 호텔의 등급에 따른 선택도 가능하게 되어 있다.

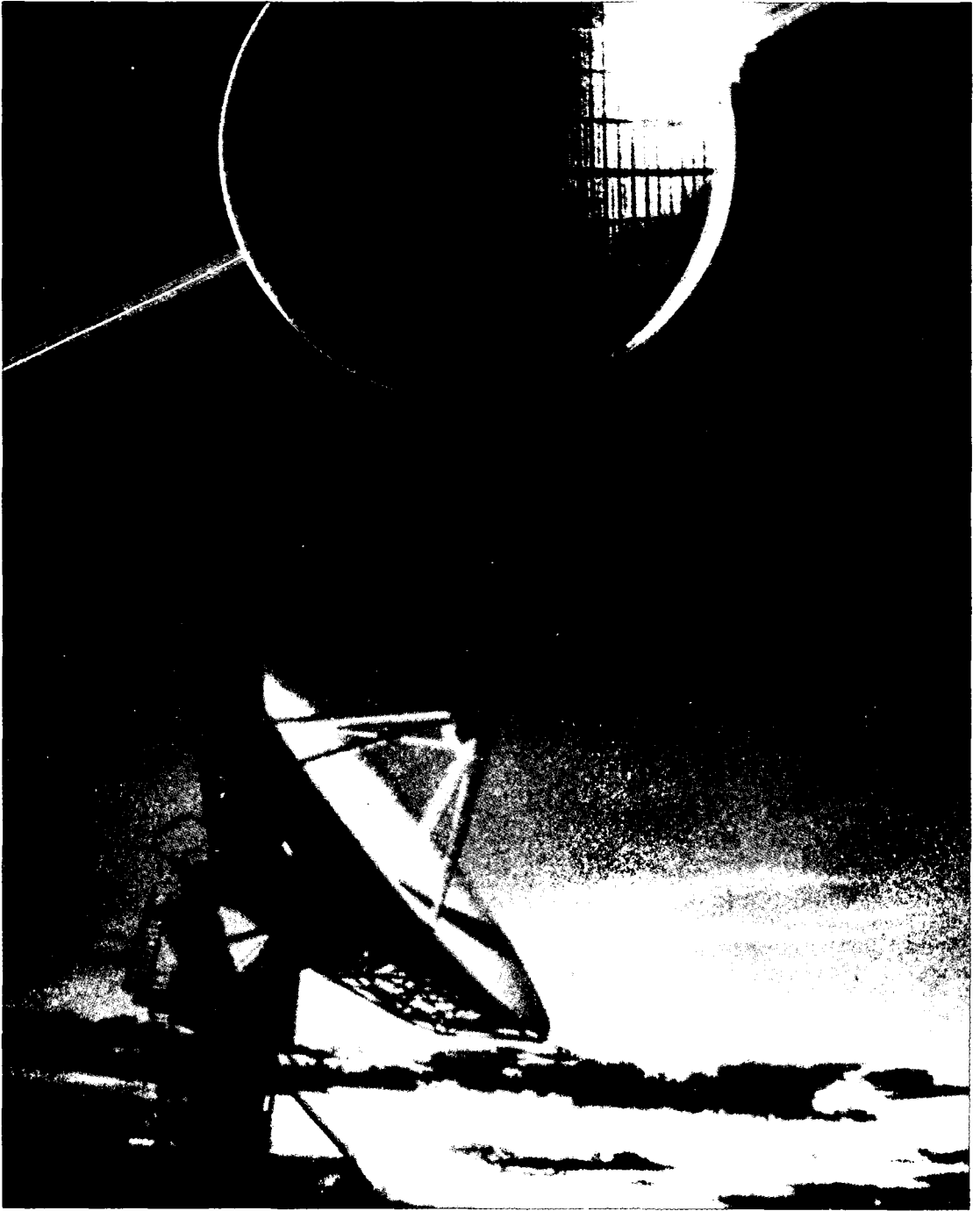
하지만 해외정보와 견줄 수 있는 DB를 개발하기에는 현재 국내의 경우 정보시장의 규모가 작은 점에서 고품질 정보개발에 어려움은 있겠지만 정보수요자의 정보요구와 부합될 수 있는 질이 좋은 DB개발 만이 정보화사회의 요구에 성공적으로 대처할 수 있겠다.

- 공간적, 시간적 요인으로 분산되어 있는 현황 정보를 통합한 정보개발
 - 이용자의 정보요구를 반영한 고품질 정보서비스 개발
- (참고:일본 DB진흥센터에서 조사한 <일반 정보이용자의 이용희망 분야>)

3.3 축적된 데이터를 관리하기 위한 국내정보검색시스템의 개발지원

전세계적으로 유명한 DOALOG는 과학, 기술, 경제, 비즈니스, 특허, 통계 등 전문분야의 정보를 포함한 약 3억건의 레코드를 수록한 방대한 데이터베이스 시스템이다. 하지만 DIALOG에서는 DIALINDEX를 사용하여 원하는 정보가 포함된 데이터베이스에 자동으로 접근시켜 주고, 검색하고자 하는 데이터베이스에 수십만건이 수록되었다고 하더라도 이용자가 원하는 정보를 불과 몇 초내에 검색할 수 있는 편리한 정보검색 서비스를 제공하고 있다.

실제로 DB개발에 있어서 정보의 축적도 중요하지만 축적된 정보를 어떻게 손쉽게 찾을 수 있도록 지원하는 것도 몹시 중요한 문제이다. 현재 국내에서 대량의 정보를 소장하고 있는 많은 기관에서 DB개발과 관련하여 안고 있는 문제점 중의 하나가 주제검색을 지원하는 국내정보검색시스템의 개발에 따른 어려움이라 하겠는데 그것은 일개 개별기관, 단체 혹은 IP가 독자적으로 문제를 해결하기에는 국내 기반 조성이 너무 취약한 실정이다. 따라서 정보검색을 위한 기반조성과



더불어 저렴하게 범용적으로 사용할 수 있는 국내정보 검색시스템의 지원이 우선되어야 한다. 이를 위해서는 구체적으로 다음사항에 대한 개발 지원이 필요하다고 본다.

○ 자연어검색을 지원하는 정보검색시스템의 개

발

- 우리말 동의어 사전의 개발
- 자동색인기법에 대한 연구투자
- 불용어 및 조사테이블의 표준화
- 일본어 표기의 표준화

또한 앞으로 전자출판으로 인한 全文정보 수록이 보편화되고 DB수량과 데이터수가 급속하게 팽창할 전망이다. 이에 따른 정보검색서비스의 개발은 더우기 시급한 일이라고 하지 않을 수 없다.

4. 맺음말

최근 국내의 정세는 국제화를 향한 피할 수 없는 대세를 피부로 실감케 하고 있다. 이미 국제화는 시작되었고 이에 따른 정보화 요구는 계속 늘어나게 될 것이다.

이제 국가적 과제가 되고 있는 정보선진국을 향한 성공적 압축성장을 위해서는 DB개발의 수량적 측면뿐만 아니라 자연스럽게 DB개발이 진척될 수 있는 기반조성과 육성책이 필요하다. 즉, 정보원처지에서의 정보공개에 대한 인식 전환, 정보를 손쉽게 DB화할 수 있는 패키지형 정보검색 시스템 개발, 정보생산성이 높은 DB의 개발 육성, IP의 대형화 유도 등이라 하겠다.

현재 국내의 DB시장은 이용자증가 등 성장세에 들어서고 있기 때문에, 보다 양질의 DB를 개발하며 정보이용 수요가 예상되는 다양한 새로운 분야를 개척함으로써 다가올 2000년 정보화 사회의 국가자원으로서 굳건한 토대를 마련할 수 있을 것이다.

참고자료 : 일반이용자의 정보이용 희망분야(일본 DB진흥센터 조사)

(1) 교육, 서적, 문화, 국제관련 등 분야

공개연구회, 각종 강습회의 일정, 요금 등 세계각국의 관습, 외국 인물정보
국가사업으로 수행되는 관민협동연구소개 단체, 대학, 그룹 등 기관 안내
고대사 관례(시대, 지역별 검색)
의약품 부작용 정보
가사지원 관계정보
기업, 단체 등 해외에서의 활동사항

내구소비재의 정가 및 시장평균 가격
각종 규격정보

(2) 취미, 오락, 레저관련 분야

호텔예약(공실)정보
관광지 시설정보(화상정보, 교통기관의 예약을 포함)
스포츠 기록, 개최예정 스케줄
스포츠 설비, 수집정보
TV스케줄 (TV드라마, 출연자 상세)
AV/Spft정보
게임정보
예술, 예능 관계 행사정보
도심정보, 전문점 정보
신차, 중고차 정보

(3) 컴퓨터 관계분야

정보제공원 정보(DATABASE관련 정보)
컴퓨터 관련 용어
컴퓨터 관계 최신 정보
제품정보(기능, 가격)
Soft의 특징
OA기기 소개
각종 부품정보
잡지 등 게재 정보

(4) 교통관계 분야

시각표(열차, 비행기, 선박, 버스, 목적지의 도착시간 부터 역산할 수 있도록)
교통수단, 도로지도
고속도로 교통정보
도로공사 정보

(5) 기타

청과물 시장가격 정보
전기제품 종합 카다로그
市, 町, 村 단위의 생활정보
알뜰시장 정보
병원정보, 병원의 대기시간 일간정보
각국의 행정사례