

공공 데이터베이스의 개발 현황에 대한 분석⁺

- 정보개발지도를 중심으로 -

김 병 록*

Analyzing the State of Public Database Development with an Information Map

Byung Rok Kim

〈요 약〉

본 연구는 국가정보화의 총괄적인 정보개발계획을 수립하기 위하여 일명 정보개발지도를 이용하여 국가의 정보개발 현황을 분석하고 국가 정보화의 방향을 제시하고자 한다. 이는 현재 국내에 개발되어 서비스를 제공중인 정보, 사업의 지원자금으로 개발을 추진중에 있는 정보, 혹은 아직 미개발되었지만 공공정보로써 개발이 가능한 공공기관의 정보자원을 조사하고 세부분야별로 분류하여 하나의 평면에 나타냄으로써 국내 공공정보 개발의 방위를 파악하고자 한다. 또한 해외의 정보개발, 서비스들을 평면에 함께 제시함으로써 향후 공공데이터베이스의 개발 과제를 파악하기 위한 기반자료로 활용될 수 있으며, 정보개발지도에 따라 년차별 개발분야 및 개발 우선순위 등을 결정될 수 있으리라 예측된다.

⁺ 본 연구는 정보통신부, 한국통신, 그리고 한국데이터베이스진흥센터의 후원과 한국외대 이주현교수의 협조로 이루어졌다.

* 초당산업대학교 산업행정학과 교수

1. 서론

1. 연구의 배경 및 목적

21세기를 향하는 세계는 개방된 정보경쟁시대로 들어가고 있다. 다가오는 세기를 대비하여 각국의 정부는 국가경쟁력의 확보를 위한 정보화촉진 및 정보서비스의 개발·보급에 관리의 초점을 두고 있다. 실로 금세기말은 정보기술에 의한 관리의 르네상스를 맞이하고 있다 하겠다 [Davenport 1993; Hammer and Champy 1994]. 정보와 정보기술에 대한 정책은 정부정책의 핵심분야가 분명하다. 정보를 보다 효율적으로 생산하고 관리하여 이를 일반 시민에게 제공하는 것은 정부조직뿐만 아니라 일반 조직에 있어서 관리의 주업무가 되고 있다. 모든 조직의 생존과 발전은 이제 정보산업의 발전과 밀접한 관계를 갖게 되었다.

데이터베이스산업은 특히 높은 부가가치를 창출하는 기술·지식 집약적 산업으로서 성장잠재력이 매우 클 뿐만 아니라 개인·공공기관·산업체 제부분의 경제·사회활동의 효율성을 높여 경제활성화 촉진 및 국민 편의를 증진하는 산업이다 [이국희 1995; DPC 1995. 7a]. 따라서 주요 선진국 등은 국가전략산업으로 육성하고 있는 실정이다. 우리나라도 최근 데이터베이스산업을 정부 차원에서 전략적으로 육성하여 국가사회 및 산업부분의 정보화를 촉진함으로써 대외 경쟁력 제고와 선진 경제를 실현코자 다양한 정책을 펼치기 시작했다 [정윤 1995; 이재홍 1995].

특히 정보통신부는 정부의 신경제 5개년 계획에 의한 정책과제 “데이터베이스 산업육성 대책(’93. 8. 19)” 수립에 따른 세부실천계획의 일환

으로 ‘공공데이터베이스 개발·보급사업’을 4년(’94~’97)개년 계획으로 추진중에 있다. 이 사업은 사업성 결여, 정보통신 이용 요금의 부담, 표준화의 미흡, 전문인력의 부족, 지적소유권의 미정립 등으로 침체되어 있는 국내 데이터베이스 산업의 활성화와 더불어 국가·사회의 정보화촉진화 및 공공부문의 정보자원공개를 촉진시켜 국가 경쟁력 강화 및 지역간 정보격차를 해소시킬 것으로 기대된다 [정보통신부 1995; 한국통신 1995]. 그러나 첫 2년간의 사업추진 중간결과 공공데이터베이스의 개발분야선정에 있어서 다음과 같은 문제점들이 대두되었다. 즉 데이터베이스 개발자에 의한 과제 제안으로 당해년도 개발과제를 신청함으로써 시장전체의 수급관계를 고려한 총괄적 정보개발계획의 부재, 비체계적인 선정으로 개발비용의 중복성 및 편협성 내재, 그리고 개발과제 수(매년 100종)의 과다로 공공데이터베이스(정부 및 공공기관 보유 정보)의 의미 희석등이다 [DPC 1994, 1995. 7b].

따라서 본 연구는 총괄적 정보개발계획을 수립하기 위한 기초작업으로서 일명 ‘정보개발지도’를 제작하여 국가의 정보개발 현황을 분석하고 앞으로의 개발과제를 밝히는데 있다. 이는 현재 국내에 개발되어 서비스를 제공중인 정보, 사업의 지원자금으로 개발을 추진중에 있는 정보, 혹은 아직 매 개발되었지만 공공정보로써 개발이 가능한 공공기관의 정보자원을 조사하여 세부 분야별로 분류하여 하나의 도면에 나타냄으로써 국내의 공공정보 개발의 방위지표를 파악할 수 있게 한다. 또 실질적으로 우리의 공공정보의 개발·서비스의 역사는 짧기 때문에 선진화된 외국의 정보개발·서비스들을 지도에 나타냄으로써 향후 공공정보개발의 방향을 파악할 수 있게 한다. 즉 이 지도는

향후 공공데이터베이스의 개발과제를 파악하기 위한 기반자료로 활용될 수 있으며, 정보개발지도에 따라 년차별 개발분야 및 개발우선순위 등이 결정될 수 있으리라 예측된다.

2. 연구 방법

본 연구의 목적은 일종의 정보개발지도를 제작하여 국내 공공데이터베이스 개발의 현황과 개발 방향을 제시하고자 하는 것이다. 따라서 본 연구는 정보개발지도에 대한 체계적 연구를 비롯하여 국내 공공정보개발의 현황을 파악하기 위한 국내 및 해외의 데이터베이스현황조사 및 국내 정보자원조사와 향후 공공정보개발사업을 위한 건의 등을 포함한다.

따라서 본 연구를 수행하기 위해서는 전문가들의 면담을 비롯하여 문헌조사, 내용분석 및 조사표에 의한 조사 등의 방법을 사용되었다.

먼저 국내 공공데이터베이스의 개발현황을 표현할 분포도로써 정보개발지도에 대한 개념적 논의를 위하여 문헌조사뿐만 아니라 정보통신부, 한국통신, 하이텔, 데이콤, 한국데이터베이스진흥센터, 그리고 학계 전문가들에 대한 면담과 5차례의 논의를 통하여 정보개발지도의 틀을 마련하였다.

다음으로 국가정보자원을 조사하기 위해서 조사표를 작성하여 중앙행정기관을 비롯하여 총 385개 공공기관들에 대하여 조사를 실시하였다. 조사표는 정보통신부장관의 협조공문과 함께 한국데이터베이스진흥센터에 의해서 약 4주(1995. 8. 21~1995. 9. 20)에 걸쳐서 실시되었다. 또한 조사표에 대한 낮은 응답.회신율을 대비하여 조사원들이 정부기관에서 발생하는 자료에 대한 내용 분석을 중심으로한 탐색조사를 하였다. 탐색조사

는 자료들이 보관되어 있는 국회도서관, 중앙도서관, 정부간행물제작소, 정부기록보관서의 자료에 대하여 실시하였다. 국내 유통 데이터베이스에 대한 조사는 정보자원조사와 함께 조사표에 의해서 조사를 실시하기도 하고, 또 한편으로는 한국데이터베이스진흥센터에서 발행하는 국내 데이터베이스목록집을 분석하였다. 해외 데이터베이스 현황 또한 해외 데이터베이스를 소개하는 문헌을 중심으로 조사하였다.

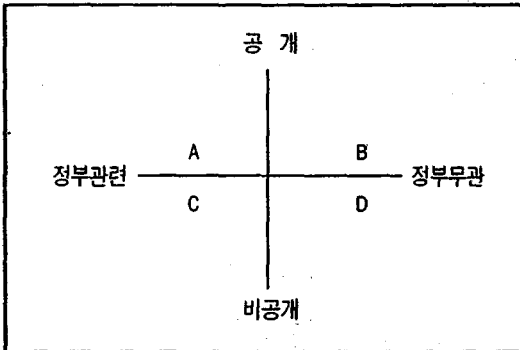
I. 공공 데이터베이스와 정보개발지도의 개념적 논의

1. 공공 데이터베이스의 개념

1) 공공 데이터베이스의 정의

데이터베이스(database)란 관련된 자료들의 수집체(a collection of related data)이다. 데이터베이스는 자료들의 출처, 현실 세계의 일과 관련성, 그리고 관심을 갖는 청중들을 갖는다[Elmasri & Navathe 1994]. 공공데이터베이스는 연구자의 관점에 따라 다양하게 논하여지고 있지만, 이에 대한 합의에는 아직 도달하지 못하고 있는 실정이다. 그러나 일반적으로 공공데이터베이스는 원자료의 공개성과 정부관련성의 특성에 따라서 데이터베이스의 개념적 영역을 살펴 볼 수 있다. 즉 유통측면에서 보아 공공데이터베이스는 산재한 원시자료를 집적하여 공중망을 통하여 공중의 이용자에게 서비스 상품 형태로 제공되는 것으로 정의되어 광의적인 의미를 갖는다. 이러한 의미의 데이터베이스는 아래 그림에서 A와 B의 영역에 속하는 것이다. 다른 한편으로 정보의 공개성여부

를 묻지않고 정부와 관련된 정보를 공공정보 또는 공공데이터베이스로 간주할 수도 있다. 보다 협의적인 의미로써 공공정보 또는 공공 데이터베이스는 정부관련 공개 정보로써 C의 영역에 속한다.



〈그림 1〉 공공 데이터베이스의 개념적 영역

2) 공공데이터베이스의 분류

공공 데이터베이스는 목적에 따라서 다양한 기준에 의해서 분류된다[고영만 외 1994]. 서비스 되는 데이터베이스에 대한 일반적인 분류기준은 주제, 기관, 제작목적, 가공정도, 표현형태, 갱신 주기, 검색방법, 서비스 주제, 서비스 형태, 이용자 범위 등이다. 그 중에서 공공 데이터베이스에 대한 가장 일반적인 분류는 주제에 의한 분류이다. 그러나 주제에 의한 분류도 다양하게 이루어지고 있는데, 가장 크게는 사용자위주와 개발자 또는 개발사업자위주로 분류되기도 한다.

대개 사용자위주의 분류는 현재 정보를 서비스 하고 있는 서비스업체에서 서비스정보에 대한 사용자의 쉬운 접근을 위해서 분류하는 것으로, 이는 전제적으로 이미 존재하는 서비스들에만 한정되는 분류이다. 따라서 계속되는 서비스의 개발에 의해서는 그 분류체계가 변화되기 마련인 것이다. 또한 서비스업체는 사용자의 쉬운 접근을 위하여

서비스에 대한 복수의 경로를 제시하기 마련이므로 이러한 분류의 기준에서는 정확한 일관성을 찾기 힘들다. 반면에 개발자 또는 개발사업자의 입장에서 분류는 현재 개발된 또는 서비스되고 있는 분야뿐만 아니라 개발가능한 분야에 대해서도 분류시에 고려하여야 한다. 이러한 경우 아직 정보 개발의 정도가 미약한 나라의 경우는 분류가 무모할 수도 있다. 하이텔, 데이콤, 그리고 미국의 Compuserve의 경우는 사용자위주의 분류라 할 것이며, 한국데이터베이스진흥센터에 의한 분류 [DPC 1995. 3.]는 상대적으로 개발사업자입장에서 있는 것이라 하겠다.

2 정보개발지도(Information Development Map)의 개념

1) 정보개발지도의 정의

지도(Map)란 지표의 일부 또는 전체를 축소시켜 각종 기호와 문자를 사용하여 평면에 그림으로 표현한 것을 의미한다[Random House 1987]. 따라서 정보개발지도(Information Development Map)란 정보개발의 일부 또는 전체를 축소시켜 각종 기호와 문자를 사용하여 평면에 표현한 것으로 정의할 수 있으며, 이를 통하여 정보개발의 현 위치와 목적지간의 방위를 파악함으로써 앞으로의 나아갈 방향들을 제시하여 주는 것이라고 하겠다.

지도에는 일반적인 내용을 종합적으로 나타내 주는 일반도와 특정한 사항만을 중심으로 선택하여 나타내주는 주제도가 있다[동아출판사 1987]. 정보개발의 일반도는 정보개발지도에 대한 필요성을 한번에 충족시켜줄 수 있도록 작성되는 지도이고 주제도는 파악하고자 하는 세부 목적에 따라 다양해 질 수 있다. 가령, 제작기관지도, 이용자지

도, 표현형태지도 등이다.

2) 정보개발지도의 필요성

정보개발지도는 다음과 같은 역할에 대한 기대에 의해서 그 필요성이 논의될 수 있다. 즉 첫째는 일종의 렌즈와 같이 사회 전체적 정보개발의 현황을 한눈에 파악할 수 있는 도구로써 정보개발 현황에 대한 모든 정보를 담고 있어야 한다. 따라서 데이터베이스개발가능분야, 개발된 분야, 그리고 개발되지 않은 부분이 동시에 파악될 수 있어야 한다. 둘째로, 체계적인 정보기획의 기초자료로써 공공정보개발추진을 위한 의사결정을 지원하기 위해서 각 분야에 대한 세세한 정보도 포함하고 있어야 한다. 정보개발지도는 또한 공공데이터베이스에 대한 외연적 (광의 및 협의의) 정의가 가능하여야 한다. 즉 현실적인 자료의 분석을 통하여 궁극적인 공공데이터베이스의 범위를 정의할 수 있다.

Ⅲ. 정보개발지도의 틀과 조사설계

1. 정보개발지도의 틀

지도를 구축하는 틀은 지도가 담고 있어야 할 내용과 각 내용의 영역에 대한 구획을 가능케하는 경계선에 의해서 결정된다. 향후 공공정보개발사업의 의사결정을 위한 정보개발지도가 담고 있어야 할 내용과 구획은 다음과 같이 제시될 수 있다.

1) 지도의 내용

정보개발지도가 담고 있어야 할 내용은 지도작성의 목적에서 추론해 낼 수 있다. 우선 지도작성의 일차적 목적은 총괄적인 국가정보개발현황을

파악하는 것이다. 따라서 정보개발지도는 국가정보자원과 개발된 데이터베이스를 내포하고 있어야 한다.

이차적으로는 공공데이터베이스 개발 우선분야를 설정할 수 있는 자료를 포함하고 있어야 한다. 따라서 공공데이터베이스 개발 현황에 대해서 비교할 수 있는 대상에 관한 자료가 있어야 한다. 이를 위해서는 비교대상으로서 해외의 선진국가의 데이터베이스 개발 현황이 필요하다. 또한 개개의 정보자원이 데이터베이스로 개발될 필요성과 시급성뿐만 아니라 자료의 성격과 형태, 그리고 주요 이용자와 이용빈도에 대한 정보도 필요하다.

다음으로 데이터베이스로 개발될 자료의 규모를 측정할 수 있는 정보가 요구된다. 따라서 정보자원의 규모를 파악할 수 있는 자료의 발생빈도 및 건당 분량 등에 대한 정보를 수록하고 있어야 한다.

2) 지도의 구획

정보개발지도의 내용의 분포를 파악하기 위한 정보개발지도의 경계선이 필요하다. 정보개발 분포비교를 위한 분포 영역의 경계선은 데이터베이스의 분류기준에서 찾을 수 있다. 본 연구에서는 지도의 경계선으로 전술하였듯이 일반적으로 가장 많이 사용되는 사용자위주의 주제별 데이터베이스 분류법을 사용한다.

사용자위주의 분류법은 전술하였듯이 많은 한계성을 갖고 있지만, 본 연구에서는 정보통신부, 한국통신, 한국PC통신, 데이콤, 한국데이터베이스진흥센터 등의 전문가들과의 면담과 5차 걸친 논의를 한 결론은 이러한 분류법이 현재의 데이터베이스 현황을 파악하고자 하는데 적합하다는 판

단이다. 이러한 분류법은 현재 정부가 추진하고 있는 '열린정부 서비스'사업에 포함되는 모든 공공정보를 수용하여 사용자가 이를 쉽게 파악하고 접근할 수 있도록 하고 있다. 또한 이 분류법은 향후 공공 정보개발사업을 시행하는 정책결정자

의 의사결정에도 요구되는 입장이라는 결론이다.

본 연구에서는 사용자입장에서 제시되어 이용되고 있는 기존의 하이텔과 데이콤 등의 분류법을 종합하여 다음 표 1과 같은 분류법을 새로 작성하여 정보개발지도의 경계선으로 이용한다.

〈표 1〉 지도 구획을 위한 사용자위주의 공공정보 분류법

대분류	중분류	중분류	중분류
1. 사회일반	1) 정치/행정 2) 통일/국제 3) 언론/인물 4) 법률 5) 종교 6) 지역 7) 당첨/공모/안내 8) 사회복지 9) 사회단체	5. 증권/금융 6. 경제/경영	1) 증권 2) 자금/환율 3) 은행 4) 투자신탁 5) 보험 1) 경제/경영일반 2) 노동 3) 기업/경영자 4) 무역
2. 생활	1) 가정/여성 2) 건강/의학 3) 상품/소비자 4) 교통정보 5) 자동차 6) 부동산 7) 날씨 8) 생활상담	7. 산업	5) 물가/시세 6) 세무 7) 경제산업 1) 산업동향/신기술 2) 공업진흥 3) 농림축수산업 4) 자원/에너지 5) 건설
3. 문화/레크리온	1) 문화/예술 2) 스포츠 3) 레저 4) 관광/여행 5) 게임/ 오락 6) 연애/ 방송 7) 모델/패션/CF 8) 만화 9) 문학 10) 영화	6) 제조업 8. 정보기술/통신 9. 과학/기술	7) 유통 8) 전기/전자 9) 금속/비금속 10) 입찰/경매 1) 컴퓨터 2) 정보통신 3) 멀티미디어 1) 과학/기술일반 2) 환경/공해 3) 규격
4. 교육/도서	1) 교육/학술 2) 유아·초등교육 3) 중등교육 4) 대학/대학원 5) 사회교육 6) 취업/자격증 7) 도서/문헌 8) 정보 9) 번역	10. 영문/해외 11. 기타	3) 규격 4) 특허/ 상표 1) 영문정보 2) 해외정보 1) 기타

2. 조사설계

본 연구를 위한 자료의 조사는 다음과 같이 국가정보자원, 국내 공공데이터베이스 현황, 그리고 해외(미국과 일본) 공공데이터베이스 현황에 대해서 이루어졌으며, 각 분야에 대한 구체적 조사 방법은 다음과 같다.

1) 국가 정보자원

국가 전체적으로 보유·관리하고 있는 정보자원의 실태를 정부기관 및 공공기관을 중심으로 파악함으로써, 공공데이터베이스의 개발과제를 파악하기 위한 기초자료로 활용하는 것이 목적이다. 따라서 정부기관 및 공공기관이 보유하고 있는 공개가능하면서 공공데이터베이스로서 가치가 있는 모든 자료를 조사하였다.

(1) 조사의 방법

국가 정보자원 조사는 조사표에 의한 조사, 그리고 정부 간행물을 제작, 보급, 또는 보관하는 기관들에 대한 탐색을 통하여 조사되었으며, 각 방법에 의해서 조사된 자료들에 대하여 조사의 중복성과 공공 데이터베이스로서의 타당성을 확인하였다.

① 조사표에 의한 조사 - 정보자원에 대한 조사표는 자료의 유형과 성격, 자료의 이용자 및 이용빈도, 자료의 내용, 자료의 규모 및 발생주기, 데이터베이스개발의 중요성과 시급성 등을 묻는 총 17개의 문항으로 구성되어 있다.

② 탐색에 의한 조사 - 본 조사는 우편으로 발송되는 조사표에 대한 낮은 회수율, 조사표 작성자의 조사에 대한 불충분한 이해 등으로 인한 조사의 불만족을 보충하기 위한 방법으로 채택되었다. 조사는 각 정부기관 및 공공기관의 간행물에

대한 조사이며, 이는 각 간행물을 보유·관리하고 있는 정부간행물제작소 자료실, 국립중앙도서관, 국회도서관, 그리고 총무처 정부기록보존소를 방문하여 각 자료를 탐색하였다.

③ 자료확인 - 정부기관이 간행한 간행물을 중심으로 한 자원조사는 조사표에 의한 조사와 중복될 수 있다. 따라서 이에 대한 서로의 대조 확인이 요구된다. 또한 탐색에 의한 조사의 결과가 조사표에 의한 조사에서 결과될 수 없는 경우도 있다. 따라서 이러한 경우는 각 정부기관에 자료에 대한 확인 절차가 요구된다. 이는 요구되는 조사표의 양식을 충족시킬 수 없는 경우에 이를 보충하기 위해서도 요구된다. 자료확인의 절차는 조사표 회수이후에 진행되었다.

(2) 조사의 대상

국가 정보자원에 대한 조사는 표 2에서 볼 수 있듯이 중앙행정기관 43개처를 비롯한 총 385개처를 대상으로 하여 이루어졌다.

(표 2) 조사 대상의 개관

분류	조사 대상의 수	분류	조사 대상의 수	총계
중앙행정기관	46개처	재정지원기관	42개처	385개처
각시도	16개처	공공법인체	40개처	
각시·도 교육청	15개처	입법·사법기관	8개처	
정부투자기관	24개처	단체 및 협회	160개처	
정부재투자기관	26개처			

2) 국내 데이터베이스 현황

국가의 가용 정보자원의 실태에 對比하여 국내에서 개발되어 유통중이거나 개발중인 공공데이터베이스의 현황을 파악하였다.

(1) 조사의 대상 - 본 조사의 대상은 국가 정보자원의 조사대상인 정부기관 및 공공기관의 385

개처를 비롯하여 국내에서 공공데이터베이스를 유통하거나 개발하는 모든 단체를 대상으로 이루어졌다.

(2) 조사의 방법 - 국내 데이터베이스현황에 대한 조사는 정부기관 및 공공기관 385개처에 대한 조사표에 의한 조사와 한국데이터베이스진흥센터에서 발간하고 있는 국내 데이터베이스 현황에 대한 문헌-1996 한국데이터베이스 목록-을 조사하는 방법으로 이루어졌다.

① 조사표에 의한 조사 - 정부기관 및 공공기관의 데이터베이스현황에 대한 조사표는 데이터베이스의 성격과 유형, 규모, 그리고 제작 및 유통 등을 묻는 총 12개 문항으로 구성되어 있다.

② 국내 데이터베이스 현황에 대한 문헌 조사 - 국내 데이터베이스 현황에 대한 문헌조사는 한국데이터베이스진흥센터에서 조사하여 발간할 예정인 "1996 한국데이터베이스 목록"을 중심으로 조사된다. 이 목록집은 국내에서 상업적 또는 공공의 목적(In-House 데이터베이스 즉 사내용 데이터베이스 제외)으로 유통되고 있거나, 개발중인 데이터베이스에 대한 현황을 조사한 결과가 수록되어 있다. "1996 한국데이터베이스 목록"은 공공데이터베이스의 현황을 데이터베이스의 내용, 데이터량, 수록기간, 제공망, 검색방식, 개발현황, 갱신주기, 그리고 제작업체 등을 중심으로 조사하였으며, 한국데이터베이스진흥센터 자신의 공공데이터베이스에 대한 분류기준인 주제별 분류기준을 사용하여 이를 분류 제시하고 있다.

3) 해외 데이터베이스 현황

국가의 가용 정보자원과 국내 공공데이터베이스의 유통 실태에 對比하여 해외, 특히 미국과 일본에서 개발되어 유통중인 데이터베이스의 현황

을 조사하였다. 선진국가의 데이터베이스의 유통 현황을 파악하여 국내 데이터베이스 유통 실태와 비교함으로써 국내 데이터베이스의 앞으로의 개발방향을 추론한다.

(1) 조사의 대상 - 본 조사의 대상은 데이터베이스산업분야의 선진국인 미국과 일본의 유통 데이터베이스를 대상으로 한다. 특히 미국의 경우에는 연방정부의 수준에서 국민에게 공개되고 있는 데이터베이스 현황에 대한 조사도 포함한다.

(2) 조사의 방법 - 미국과 일본의 공공 유통 데이터베이스의 현황에 대해서는 범세계적으로 데이터베이스 현황을 조사한 Gale Directory of Database(1995)을 중심으로 조사를 하며, 미국 연방정부에서 서비스하고 있는 데이터베이스 현황에 대해서는 The Federal Database Finder를 중심으로 조사를 하였다.

해외 데이터베이스 현황에 대한 조사는 데이터베이스의 제작목적, 데이터베이스의 갱신주기, 데이터베이스의 표현형태, 데이터베이스의 사용언어, 그리고 데이터베이스의 국적을 중심으로 이루어졌다.

① Gale Directory of Database에 대한 조사 - Gale Directory of Database는 Gale Research Inc.에 의해서 1년에 두차례 발간되는 것으로, 데이터베이스현황에 대한 범세계적 조사의 결과이다. 각 데이터베이스에 대해서 데이터베이스 제작, 담당, 유형, 내용, 분류, 언어, 서비스개시일, 수록기간, 갱신주기, 유통망 등을 수록하고 있다.

② The Federal Database Finder에 대한 조사 - The Federal Database Finder는 미 연방정부로부터 자유로이 이용가능한 데이터베이스와 자료 화일에 대해서 Information USA, Inc.가 비정기적으로 발간하는 목록집이다. 이 목록집은

미국 행정부의 14개 부처를 비롯하여 독립기관들, 사법기관들, 입법기관들을 포함하여 총 63개처에서 이용가능한 데이터베이스의 목록을 수록하고 있다.

Ⅳ. 정보자원 및 공공데이터베이스의 현황

1. 국가 정보자원의 현황

국가정보자원에 대한 현황은 분석은 385개의 공공기관에 조사표를 보내어 회신된 96개기관의 265개의 정보자원과 본 연구의 조사원이 정부간행 자료들을 탐색하여 수집한 238개의 자원을 포함하는 총503개(자료확인 과정에서 중복된 자원이 128개)를 대상으로 분석한다.

조사원이 정부간행 자료들을 중심으로 가용한 정보자원으로 판단할 때에는 다음의 원칙을 준수하였다. 첫째, 비롯 자료를 수록하고 있는 문건이 많을 지라도 그것이 하나의 데이터베이스로 개발하는 것이 타당하다고 판단되는 경우, 이를 하나의 정보자원으로 간주한다. 가령, 매주기별로 간행되는 통계자료집은 비록 그 규모가 막대할지라도 하나의 가용자원으로 간주하였다. 둘째, 비록 한 자료집일지라도 그 내용에 있어서 다수의 가용자원을 포함하고 있으며, 자료의 내용에 따라서 가용자원을 구분하였다. 가령, 원자력백서와 같은 하나의 자료집에는 원자력에대한 기술정보, 원전 피해사례 등의 다수의 가용자원을 갖을 수 있다. 셋째, 조사대상기관에서 회신한 자원과 조사원이 탐색한 자원이 중복되는 경우, 회신된 자료의 내용을 우선적으로 고려한다. 본 연구과정에서 중복

된 자원들은 전술하였듯이 128개였다.

1) 자료의 유형과 성격별 현황

자료의 유형과 성격별 현황은 각 자료의 유형과 성격에 대한 빈도분포만을 파악한다.

(1) 자료의 관리형태에 대한 현황 - 국가정보자원은 일반적으로 책자의 형태로 가장 많이 관리되고 있다. 특히 데이터베이스의 형태로 관리되는 자원도 있지만, 이는 공중망에 유통되기보다는 각 기관내부에서 사용되고 있는 것으로 이를 공공데이터베이스화하는 것이 요구되기도 한다. 또한 본 연구에서 자료에 대한 탐색이 정부기관이 간행한 책자를 중심으로 이루어져 있기 때문에 여기서 관리형태로 책자의 형태가 많이 나왔다.

또한 자료들은 반드시 한가지의 형태로 관리되기 보다는 다양한 형태로 관리되기도 한다. 가령 우선적으로는 책자형태로 관리되지만, 그 양이 증대됨에 따라 마이크로필름에 옮겨서 관리되기도 한다.

(2) 자료의 유형에 대한 현황 - 자료의 유형은 통계부분이 가장 많은 비중을 차지하고 있다. 그리고 많은 자료들을 보고서나 백서 및 연감에서 발견할 수 있다.

(3) 자료의 표현형태에 대한 현황 - 수집된 정보자원의 표현형태로는 문자에 의해서 표현되는 서지와 전문 형태의 자료가 가장 많고 다음으로는 수치 및 기호에 의한 표현되는 자료들이 많이 있다.

(4) 자료의 목적에 대한 현황 - 정보자원의 주요 목적으로는 국민에 대한 공공정보 제공차원에서 작성되는 경우가 가장 많으며, 다음으로는 정부기관의 내부사용용, 자료의 기록의 차원 순으로 나타나고 있다.

〈표 3〉 자료의 유형과 성격별 현황

자료 관리형태	책자/종이	컴퓨터 파일	마이크로 필름	광판	DB	기타				
현황	430	47	6	6	66	9				
자료 유형	편람	통계	책자 및 연감	보고서	법률	공보	연설	특허	기타	
현황	23	93	31	44	7	6	·	11	86	
자료 표현형태	사지	전문	사전	수치 및 기호	화상	음성	멀티 미디어	기타		
현황	160	37	5	67	12	3	3	4		
자료 표현형태	홍보	교육	기록	공공정보 제 공	내부사용	기타				
현황	41	28	42	157	92	46				

복수 응답된 것임

2) 자료의 주요 이용자와 이용빈도별 현황

자료의 주요 이용자와 이용빈도별 현황에 대해서 각 항목에 대해서 관련 자료의 빈도분포만을 제시한다.

(1) 자료의 주요 이용자에 대한 현황 - 자료의 주요 이용자에 대한 현황을 볼 때, 각 자료가 한 가지 유형의 이용자집단에 의해서 요구되기보다는 다수의 집단들에 의해서 요구되며, 자료중에는 일반대중에 대한 정보제공뿐만 아니라 자체 또는 타 기관에 의해서 이용되는 자료도 많이 있음을 알 수 있다.

(2) 자료의 이용빈도에 대한 현황 - 자료의 이용빈도는 월 평균적으로 조사되었으며, 최대빈도와 최소빈도를 조사한 후 이들 사이의 간격을 5개의 간격으로 구분하여 현황을 보면 다음 표와 같다. 자료중에는 월 300,000회 정도가 이용되는 자

〈표 4〉 자료의 주요 이용자와 이용빈도별 현황

자료 관리형태	기업	학자	학생	시민	내부	다 른 정부기관	기타	복수 응답
현황	130	129	73	70	156	120	34	
자료 유형	~100	~500	~1000	~10000	~300000	기타	계	
현황	53	38	16	23	12	29	171	

료가 있음을 알 수 있다. 자료 이용빈도의 중앙치는 500회부터 1000회까지로 볼 수 있다.

3) 자료의 규모별 현황

자료의 규모별 현황에 대해서 각 항목에 대해서 관련 자료의 빈도분포만을 일정한 리커트를 사용하여 제시한다.

(1) 자료의 현재 보존량에 대한 현황 - 자료의 보존량은 자료관리의 형태에 따라서 건수와 Byte 수로 조사되었다. 각 보존량은 최대치와 최소치를 조사한 후 이들 사이의 간격을 5개의 간격으로 구분하여 현황을 보면 아래의 표들과 같다.

(2) 자료의 발생빈도에 대한 현황 - 자료의 발생빈도를 보면 아래 표와 같다. 연간단위가 가장 많으며, 다음으로 월간단위, 매일별 순으로 나타난다.

(3) 자료의 건당 분량에 대한 현황 - 자료의 건당 분량은 자료관리의 형태에 따라서 쪽수와 Byte수로 조사되었다. 각 자료의 건당 분량은 최대치와 최소치를 조사한 후 이들 사이의 간격을 5개의 간격으로 구분하여 현황을 보면 아래 표들과 같다. 즉 대부분의 자료의 건당 분량은 400쪽이하이거나 1000byte이하임을 알 수 있다.

(4) 자료의 보유기간 - 자료의 보유기간은 최대치와 최소치를 조사한 후 이들 사이의 간격을 5개의 간격으로 구분하여 현황을 보면 다음 표와 같다. 현 자료의 많은 부분이 30년 이하의 기간동안 보유되어오고 있다. 80년 이상 보유되고 있는 정보자원은 없는 것으로 나타났다.

(5) 자료의 보존기간 - 조사대상의 기관에서 각 자료에 대해서 설정해둔 자료의 보존기간은 최대치와 최소치를 조사한 후 이들 사이의 간격을 5개의 간격으로 구분하여 현황은 다음 표와 같다. 여

기에서 기타는 영구와 준영구를 포함하고 있으며, 많은 자료들이 영구기간을 갖고 있다.

〈표 5〉 자료의 규모별 현황

자료 보존량	건 현황	~1만	~10만	~100만	~1000만	~1억건만	계	
	Byte	123	43	21	23	4	214	
	현황	~1M	10M	~100M	~200M	~20G	계	
		2	2	5	1	3	13	
발생빈도		매일	매주	매월	분기	매년	비정기	기타
		35	18	92	27	210	93	25
자료 건당 분량	건 현황	~400	~800	~1200	~2000	~2만	계	
	Byte	304	98	24	7	6	439	
	현황	~500	~1000	~2000	~3000	~4M	계	
		15	10	3	2	4	34	
보유기간		~17년	~34년	~51년	~64년	~80년	계	
		102	40	15	1	1	159	
보존기간		~6년	~12년	~18년	~24년	~50년	기타	계
		48	16	5	4	115	188	

4) 데이터베이스 개발의 필요성과 시급성별 현황

데이터베이스개발의 필요성과 시급성에 대한 현황은 조사양식에 사용된 5점 리커트에 따라서 작성된 경우에 한해서 현황을 제시한다.

(1) 데이터베이스 개발의 필요성에 대한 현황 - 정보개발의 필요성은 정보를 보유하고 있는 정부기관에서 판단한 것으로 아래표에 나타나 있듯이, 많은 자료들이 공공데이터베이스로써 개발될 필요가 있다고 주장하고 있다. 물론 개발가치가 없는 자료들도 나타나고 있다.

(2) 데이터베이스 개발의 시급성에 대한 현황 - 아래 표에서 볼 수 있는 자원들에 대한 정보개발의 시급성 또한 정보를 보유하고 있는 정부기관에서 판단한 것으로 개발이 시급한 자원이 127개로 많은 부분이 개발이 요구되고 있다

2. 국내 데이터베이스 현황

공공 데이터베이스활성화 정책에 힘입어, 국내

〈표 6〉 데이터베이스 개발의 필요성과 시급성별 현황

개발의 중요성	매우 필요하다	필요하다	보통이다	필요하지 않다	전혀 필요하지 않다	계
	91	74	33	10	11	219
개발의 시급성	매우 시급하다	시급하다	보통이다	시급하지 않다	전혀시급하지 않다	계
	70	57	64	10	10	211

공공데이터베이스 제작의 활성화를 촉진하고 있으며, 이를 통해 많은 데이터베이스의 생성과 활용이 이루어지고 있는 초기단계의 입장에 있다.

이런 측면에서 데이터베이스의 이용도 활성화와 이용자에 대한 데이터베이스제공에 대한 서비스 측면 등을 중시하고 있는 실정이다. 특히 데이터베이스의 기술적인 측면, 즉 멀티미디어의 활성화와 이미지, 화상 등을 통한 데이터베이스의 기술적인 측면은 이용자의 호응을 쉽게 받을 수 있는 이점을 갖고 있으므로 중요시 고려시 되어야 한다. 하지만 실질적으로 데이터베이스를 제작하는 목적은 정확한 양질의 데이터를 획득할 수 있도록 사용자에게 부가가치가 높은 데이터베이스를 만들어 제공하는데 가장 큰 목적이 존재한다. 이런 측면에서 국내데이터베이스는 기술적인 면과 동시에 정보의 양이 부족하다는 결점을 보강하기 위한 정보자원 획득의 측면 역시 중요하게 고려되어야 한다.

이런 측면에서 국내데이터베이스는 기술적인 면과 동시에 정보의 양이 부족하다는 결점을 보강하기 위한 정보자원 획득의 측면 역시 중요하게 고려되어야 한다.

1) 국내 데이터베이스의 유형과 성격별 현황

데이터베이스의 유형과 성격별 현황은 데이터베이스의 표현형태와 사용언어에 대한 빈도분포만을 파악한다.

(1) 데이터베이스의 표현형태에 대한 현황 - 국내 데이터베이스의 두드러진 표현형태는 다수의 데이터베이스가 멀티미디어 표현형태를 갖추고 있는 것이며, 이런 현상은 사용자의 데이터베이스 활성화에 큰 도움을 줄 수 있을 것이다. 하지만 데이터베이스의 내용을 충실히 한다는 입장에서 여전히 텍스트 형식의 데이터베이스화도 중요시되고 있는 입장이다.

(2) 데이터베이스의 사용언어 - 현재 제작중인 국내 데이터베이스의 대부분은 한글을 채택하고 있으며, 그 이유는 데이터베이스제작의 가장 큰 우선 순위를 국민 모두에게 대국민공공 정보제공과 서비스에 제공하는데 가장 큰 우선 순위를 부여하고 있기 때문이다. 하지만 미래의 데이터베이스 활용 측면을 고려하는 입장에서는 좀 더 많은 국내 데이터베이스를 상품화한 영문 데이터베이스도 필요시 되어야 한다. 일부 데이터베이스의 한글과 영문 혼합체계를 채택하여 다른 정보기관에 정보제공을 고려하는 경우의 입장도 채택하고 있다. 또 다른 측면에서는 다른 정부의 정보자원을 이용하는 입장에서도 영문을 채택하는 경우도 존재할 수 있게 된다. 하지만 순수하게 영문으로만 구성된 데이터베이스의 수는 소수이며 이런 경우의 데이터베이스는 정보제공원이 다른정부기관인 경우가 많다. 데이터베이스의 상품화하여 다른정부기관에 정보제공을 하는 측면도 고려 대상에 포함되고 있다.

(표 7) 유명과 성격별 현황

표현형태	서지	수치 및 기호	전문	기타	계
	211	33	777	245	1266
사용언어	한글	영어	한글+영어	기타	계
	747	53	442	24	

2) 데이터베이스의 규모별 현황

(1) 데이터베이스 수록기간 - 국내 데이터베이스현황에 대한 조사에 따르면, 국내 데이터베이스의 자료수록은 1940년대의 자료가 가장 오랜 것이며(5개), 대부분의 데이터베이스가 1990년대의 자료를 수록하고 있다. 여기서 기타는 그 수록의 시기가 미확인 된 경우이다.

특히 국내 데이터베이스는 데이터베이스제작이 늦은 것만큼 데이터베이스화한 정보의 양부 족과 데이터수록기간의 단기간으로 통한 데이터베이스의 수록 내용이 불충분한 약점을 가지고 있는 입장므로 꾸준한 정보자원 개발이 필요시 되는 실정이다. 결국 데이터베이스의 생명은 많은 양의 데이터를 최적 화시켜 이용자의 요구사항을 최대한 만족시켜 주는 데 있으므로 국내 데이터베이스의 질과 양적인 면을 강화하기 위한 전략적인 면에서 국내 데이터베이스의 정보축적이 우선적으로 필요한 실정에 있다.

(2) 데이터베이스의 갱신주기 -- 국내 데이터베이스의 갱신주기는 비정기적으로 이루어지는 경우가 가장 많으며(560), 다음으로는 매일 갱신이 이루어지는 경우(218), 월간(156), 연간(121), 주간(87), 분기별(41) 순이며, 갱신이 이루어지지 않는 데이터베이스의 경우도 18개 있다.

실질적으로 데이터베이스를 갱신하는 유형은 여러 유형을 갖출 수 있지만, 국내 데이터베이스의 경우는 대 다수가 매일 또는 비정기적 갱신주기 유형을 채택하고 있다. 몇 가지 예로써 데이터베이스의 내용이 언론과 잡지에 관련된 내용(뉴스, 신문, 주식시세)들은 대다수 분야의 데이터베이스는 매일 또는 주기적(주간, 월간)으로 갱신하는 경우가 되고 반면에 갱신하지 않는 데이터베이스 형태는 기록이나 보존을 중시하는 경우로써 실질

〈표 9〉 국내 데이터베이스의 규모별 현황

수행기간	1990년이전		1960년대	1970년대	1980년대	1990년대	기타	계	
	12	8	12	94	830	310	1266		
강신주기	매일	주간	월간	분기별	년간	비정기	강신없음	기타	계
	218	87	156	41	121	560	18	65	1266

적으로는 데이터베이스자체의 내용을 변경하지 않는 경우들이 대부분이 해당된다.

3) 제작체제에 대한 현황

국내데이터베이스 제작체제는 자체의 데이터베이스활용도 활성화를 위해 자체제작을 통하여 하는 경우가 대 다수이며, 이런 경우는 데이터베이스의 관리측면에서 자체데이터베이스유지보수를 진행하는데 있어 자체 내에서 계속 관리하는 경우며, 반면에 전문업체에 의뢰하는 경우는 데이터베이스의 자체제작이 어려운 상황에 있거나 데이터베이스제작에 실질적으로 관여할 수 없는 경우로 이런 경우는 데이터베이스제작 이후 데이터베이스의 유지보수 관리면 에서 약간의 문제점을 도출할 수 있다. 또 다른 경우로써는 자체의 데이터베이스제작능력을 고려하여 전문업체에 고려하는 것으로 자체의 정보 활용도 활성화와 데이터베이스제작 이후 데이터베이스의 관리측면을 고려해야 공동제작체제를 구성하는 방식을 채택하고 있는 경우이다.

〈표 10〉 국내 데이터베이스의 제작체제 현황

분류	자체	전문업체	자체+전문	기타	계
현황(%)	1050	103	112	1	1266

4) 데이터베이스사용의 유료화 현황

현재 국내공공데이터베이스의 사용료에 대한

현황은 무료와 유료의 각 비율 차가 거의 비슷한 상태이지만 대부분의 무료데이터베이스의 경우는 데이터베이스의 유지보수 관리측면에서 데이터베이스의 제작초기에는 무료화 하가

일정기간이 경과한 후에는 비용을 부담시키는 방식을 채택하고 있다.

〈표 11〉 국내데이터베이스의 유료화 현황

분류	무료	유료	기타	계
현황	558	551	157	1266

3. 해외 데이터베이스의 현황

해외의 데이터베이스에 대해서 미국의 Gale Research Inc.가 조사한 결과에 의하면, 세계의 데이터베이스산업은 규모에서 데이터베이스, 데이터베이스엔트리, 데이터베이스 제작기관, 데이터베이스 유통기관 중의 양적 증가를 보이고 있다. 즉 1975년부터 1994년까지 19년동안의 추이를 보면 데이터베이스 수는 현재 8,776개로 29배의 성장을 데이터베이스엔트리는 7,979개로 약 27배, 데이터베이스제작기관 수는 2,778개처로 14배, 유통기관 수는 1,691개처로 16배의 성장을 하였다. 또한 같은 기간동안 데이터베이스 레코드 수는 80년대 중반부터 시작된 상업용 비즈니스 데이터베이스들의 증가로 약 122배 증가하였다.

특히 Gale Directory를 통해서 본 해외데이터베이스의 추세는 데이터베이스의 수적 증가와 마찬가지로 데이터베이스제작에 대한 기술적 측면이 많이 고려되어지고 있다. 데이터베이스의 멀티미디어와 이미지,화상등이 차츰 중요시되고 있는

입장이다. 하지만 실질적으로 많은 데이터베이스가 여전히 텍스트형식으로 제작되고 있으며, 이런 가장 큰 이유는 데이터베이스의 제작목적 자체 갖는 근본적인 중요성은 많은 양질의 정보를 축적하여 사용자에게 제공하는 것에 더 우선 순위를 주고 있기 때문이다.

1) 해외 데이터베이스 제작목적별 현황

대부분의 해외데이터베이스의 제작목적은 대국민공공 정보제공과 타 기관 정보제공(다른 국가)을 주된 목적으로 제작되고 있는 입장이다. 국내 데이터베이스 제작목적이 대국민공공 정보제공에 우선 순위를 준다면 외국의 경우는 자체데이터베이스를 상품화하여 다른정부기관에 제공함으로써 데이터베이스에 대한 질적 강화를 증시하여 마케팅성화에 많은 관심을 보여주고 있다.

<표 12> 해외데이터베이스 제작목적 현황

분류	홍보	교육	기록	대국민 정보제공	내부 사용	타 기관 정보제공	기타	계
현황	32	25	22	609	97	325		1110

2) 해외 데이터베이스의 표현형태에 대한 현황

해외데이터베이스의 표현형태에 대한 비율은 텍스트형식이 대부분을 차지하고 있으며 화상과 멀티미디어의 경우는 중요성에 대한 인식과 관심은 많지만 실제적으로 데이터베이스의 표현형태로 채택된 것이 서비스중인 것은 별로 없는 실정이다. 많은 데이터베이스의 경우가 중복하여 표현형태를 사용하고 있다. 그렇지만 멀티미디어에 대한 활성화는 짧은 시간 내에 멀티미디어기술의 발전과 데이터베이스의 가치인식을 통해 발전추세를 보이게 될 것이다.

<표 13> 해외데이터베이스의 표현형태 현황

분류	서지	전문	시전	수치 및 기호	화상	멀티 미디어	기타	중복 응답된 것임
현황	420	1781	429	787	7	.	3424	

3) 해외 데이터베이스의 사용언어

미국데이터베이스의 대부분은 자체(미국)의 정보원을 사용하여 제작한 경우와 다른 나라의 정보원을 데이터베이스화한 경우 이 두 가지로 분류할 수 있으며 이들 대부분은 데이터베이스제작을 영어로 작성한 경우로 다른 나라에 정보제공 측면을 역시 중요시하므로써 데이터베이스의 원활하게 다른정부기관에 제공하는데도 큰 부과가치를 부여하고 있다.

<표 14> 해외데이터베이스의 사용언어 현황

언어	영어	기타	계
현황	2971	119	3090

4) 해외 데이터베이스의 규모별 현황

데이터베이스의 규모별 현황에 관련되는 각 항목에 대해서 관련 자료의 빈도분포만을 일정한 리커트를 사용하여 제시한다. 이에 대한 사항은 정보개발의 일반도에서도 참조할 수 있다.

(1) 해외 데이터베이스의 수록기간 - 데이터베이스의 수록기간은 연대별 간격을 5개의 간격으로 구분하여 현황을 파악한다. 국내 데이터베이스의 데이터수록기간이 짧다는 약점에 비해 해외데이터베이스의 데이터수록내용은 연대별로 꾸준하게 정보 내용을 축적해 왔다. 특히 1980에서 90년 중반까지 동안 축적된 데이터의 데이터베이스양은 두드러지게 증가추세를 보이고 있다.

(2) 해외 데이터베이스의 갱신주기 - 대부분의

〈표 15〉 해외 데이터베이스의 규모별 현황

수록기간	1990년이전		1960년대	1970년대	1980년대	1990년대	계	
	155	115	272	901	1195	2737		
갱신주기	매일	주간	월간	분기별	년간	비정기	갱신없음	계
	647	680	620	160	160	52	128	2547

데이터베이스갱신주기는 매일, 주간, 월간 형태로 갱신이 되어지고 있으며, 그외 단지 소수의 데이터베이스가 비정기적인 형태로 갱신되고 있으며, 비정기적인 형태는 국내데이터베이스에 비해 아주 작은 부분을 차지하고 있으며, 이런 측면을 통해 해외데이터베이스는 데이터베이스의 정보자원 내용을 꾸준히 유지보수하는 입장을 채택하고 있다.

V. 정보개발지도 및 개발현황 분석

1. 정보개발지도

전술한 연구의 목적에 입각하여 국내 공공정보의 데이터베이스화를 추진하기 위한 정보개발현황을 분석하고자 조사된 국내 유통 데이터베이스 현황, 국가정보자원 현황, 그리고 해외 데이터베이스 현황에 대한 조사를 수록하는 정보개발지도는 정보자원에 대한 조사결과 503개, 국내의 유통 데이터베이스 1266개, 그리고 해외 데이터베이스 3090개를 표 1에 제시된 사용자 위주 공공정보의 분류법에 의하여 수록하여야 한다. 이러한 지도가 본 연구의 목적에 보다 부합하기 위해서는 조사된 모든 자료가 한 눈에 파악할 수 있도록 보다 평면에 기록될 필요성이 있다. 그러나 막대한 자료를 한 평면에 옮기는 것은 쉽지 않다. 따라서 표 16

과 같은 분포표를 이용함으로써 정보개발현황을 우선적으로 파악하게 할 수 있으며, 보다 구체적인 정보는 각 자료에 대한 정보를 담고 있는 개별 쉬트(sheet)를 활용하는 것이 바람직 하다고 판단한다. 따

라서 본 논문에서는 각 분야별 정보개발의 현황을 분포표에 의해서만 제시하고 이를 중심으로 정보개발의 현황을 파악하고 앞으로의 개발과제를 제시하고자 한다.

2. 정보개발 현황분석

1) 정보개발의 일반적 현황

정보개발의 일반적 현황으로는 기존의 정보개발 상황에서 가장 중점적으로 이루어진 분야들을 국내.외의 경우에서 순서대로 살펴보고, 그리고 개발가능한 정보자원이 풍부한 분야를 살펴본다.

(1) 정보개발의 대분류별 현황 - 정보개발 현황에 대한 분포표에 따르면, 기존의 정보개발의 중점분야에 있어서 해외와 국내간에 차이가 있음을 볼 수 있다. 해외의 경우에는 사회일반분야(613개)에 가장 많은 데이터베이스를 보유하고 있고, 다음으로 경제/경영분야(503개), 산업분야(359개), 교육/도서분야(331개), 생활분야(322개), 과학/기술분야(314개) 순으로 데이터베이스를 보유하고 있다. 반면에 국내의 경우에는 교육/도서분야(234개)에 데이터베이스가 가장 많이 개발되어 있고, 다음으로 경제/경영분야(192개), 생활분야(174개), 사회일반분야(166개) 순으로 정보가 개발되어 유통되고 있다.

해외의 경우나 국내의 경우도 경제/경영분야 그리고 사회일반분야에 정보개발이 중점적으로이

<표 16> 정보개발연황 분표표

대분류	중분류	정보 자원	국내 DB	해외 DB	대분류	중분류	정보 자원	국내 DB	해외 DB
1. 사회 일반	① 정치/행정	31	30	142	6. 증권 /금융	① 증권	1	33	83
	② 통일/국제	7	6	9		② 자금/환율	2	14	62
	③ 언론/인물	3	43	106		③ 은행	2	37	34
	④ 법률	11	18	175		④ 투자신탁	0	1	17
	⑤ 종교	0	4	8		⑤ 보험	3	5	15
	⑥ 지역	86	50	139	계	8	90	211	
	⑦ 당첨/공모/안내	0	5	2	6. 경제 /경영	① 경제/경영일반	14	35	182
	⑧ 사회복지	15	8	28		② 노동	18	9	71
	⑨ 사회단체	0	2	4		③ 기업/경영자	11	71	155
	계	153	166	613		④ 무역	32	48	40
2. 생활	① 가정/여성	6	29	21		⑤ 물가/시세	8	13	22
	② 건강/의학	8	39	143		⑥ 세무	2	8	33
	③ 상품/소비자	2	45	66	⑦ 경제산업	0	8	0	
	④ 교통정보	13	8	45	계	85	192	503	
	⑤ 자동차	1	7	14	7. 산업	① 산업동향/신기술	6	23	92
	⑥ 부동산	1	33	23		② 공업진흥	2	11	0
	⑦ 날씨	4	6	6		③ 농림축산업	39	30	42
	⑧ 생활상담	1	7	4		④ 자원/에너지	31	13	126
	계	36	174	322		⑤ 건설	5	15	11
3. 문화 /레포츠	① 문화/예술	12	41	21		⑥ 제조업	5	15	33
	② 스포츠	2	12	17		⑦ 유통	9	5	17
	③ 레저	1	2	7		⑧ 전기/전자	2	4	15
	④ 관광/여행	4	23	24		⑨ 금속/비금속	0	6	17
	⑤ 게임/오락	0	7	9		⑩ 입찰/경매	0	7	6
	⑥ 연예/방송	0	18	25	계	99	129	359	
	⑦ 모델/패션/CF	2	4	0	8. 정보 /통신	① 컴퓨터	0	29	108
	⑧ 만화	0	5	2		② 정보통신	7	21	119
	⑨ 문학	1	0	5		③ 멀티미디어	0	3	4
	⑩ 영화	2	2	5		계	7	53	231
계	24	114	115	9. 과학 /기술	① 과학/기술일반	12	14	110	
4. 교육 /도서	① 교육/학술	16	37		93	② 환경/공해	8	4	79
	② 유아/초등교육	0	2		3	③ 규격	1	14	28
	③ 중등교육	0	1		4	④ 특허/상표	13	14	36
	④ 대학/대학원	0	6		5	⑤ 첨단과학	0	0	61
	⑤ 사회 교육	0	8	22	계	34	46	314	
	⑥ 취업/자격증	1	31	20	10. 일반 /사회	① 영문정보	1	1	0
	⑦ 도서/문헌	27	115	155		② 해외정보	10	43	91
	⑧ 정보	2	34	29		계	11	44	91
	⑨ 번역	0	0	0	11. 기타	① 기타	0	24	0
계	46	234	331	계		0	24	0	
					총 계				
					503 1266 3090				

루어진 것을 볼 수 있지만, 국내에는 교육/도서분야가 상대적으로 많은 정보개발현황을 보이는 반면에 해외의 경우에는 국가의 경제력이나 산업의 발전에 영향을 줄 수 있는 과학기술분야의 정보개발이 많이 이루어져 있음을 알 수 있다. 현재 파악된 국내.외간 데이터베이스의 상대적 비율(전체적 비율은 2.44임)에 의하면 그 차이가 가장 적은 분야는 문화/레포트분야(1.01), 교육/도서(1.41), 생활분야(1.85), 증권/금융분야(2.34) 순이다.

또한 정보자원에 대한 현황은 정부기관과 공공기관에 제한되어 최광의 공공 데이터베이스 개발을 위한 정보자원의 집중분야를 논의하기는 힘들다. 그러나 공공기관의 정보자원의 현황을 보면, 사회일반분야(153개)가 정보자원이 가장 많고, 다음으로 산업분야(99개), 경제/경영분야(85개) 순이다.

(2) 정보개발의 중분류별 현황 - 정보개발의 현황을 중분류별로 볼 때, 해외의 경우에는 경제/경영일반(182개), 법률(175개), 기업/경영자(155개), 도서/문헌(155개), 건강/의학(143개), 정치/행정(142개), 지역(133개), 자원/에너지(126개) 정보통신(119개), 그리고 과학/기술일반(110개) 등의 순으로 정보가 개발되어 있고, 국내의 경우는 도서/문헌(115개), 기업/경영자(71개), 지역(50개), 무역(48개), 상품/소비자(45개), 언론/인물(43개), 해외정보(43개), 문화/예술(41개), 건강/의학(39개), 그리고 교육/학술(37개) 등의 순이다. 한편으로 국내 정부기관.공공기관 및 공공단체가 보유하고 있는 정보자원은 지역(86개), 농림축산업(39개), 무역(32개), 정치/행정(31개), 자원/에너지(31개), 도서/문헌(27개), 노동(18개), 교육/학술(16개), 그리고 사회복지(15개) 등의 순으로 많은 것으로 나타나고 있다. 다음에

서 각 분야별로 그 일반적 특징을 살펴본다.

먼저 사회일반분야는 국내.외에서 정보개발이 많이 이루어진 분야이다. 그러나 국내의 경우에는 현존하는 정보자원에 비하여 정보개발이 적게 이루어진 부분으로도 볼 수 있다. 특히 지역분야, 사회복지분야, 통일/국제분야, 그리고 정치/행정분야의 경우에는 파악된 정보자원보다 유통되고 있는 데이터베이스의 수가 적은 실정이다.

생활·문화/레포트·교육/도서·증권/금융·정보통신/기술분야는 정부기관이나 공공기관에서 소유하고 있는 정보자원이 상대적으로 빈약하여 현재 개발된 데이터베이스들과의 비교가 힘든 분야이다. 그러나 생활분야, 문화/레포트분야, 그리고 교육/도서분야는 국외와의 비교에서 전반적으로 많은 정보가 개발되었음을 알 수 있다.

경제/경영분야는 국가의 경쟁력이나 경제발전에 주요한 분야로서 국가적 차원에서 정보개발이 요구되는 분야이다. 그러나 이 분야의 주요 정보개발은 사적분야에서 많이 이루어지고 있으며, 사실 정부기관이나 공공기관에서 공개되는 정보자원도 많이 생산되지 않고 있는 것 같다. 그러나 해외의 경우에는 정보개발의 중점분야중의 하나이다.

경제/경영분야에서는 기업이나 경영자에 관련된 정보, 경제/경영일반, 그리고 무역 등에 대한 정보가 많이 개발되어 있으며, 정보자원도 비록 적은 수이지만 그러한 분야에 집중되어 있다. 외국의 경우에는 이외에도 노동분야에 대한 정보가 상당히 개발되어 있음을 알 수 있다.

산업분야는 정보가 많이 개발된 분야이기도 하지만 또한 개발되어질 정보자원이 많은 분야중의 하나이다. 특히 농림축산업분야, 자원/에너지분야, 그리고 유통분야는 개발되어진 데이터베이스보다

는 정보자원이 풍부한 분야이다. 해외의 경우에는 이분야에서 대부분의 데이터베이스가 산업동향/신기술분야와 자원/에너지분야에 집중되어 있다는 것도 주목하여야 할 부분이다.

과학/기술분야는 해외 데이터베이스에 비해서 상대적으로 국내 데이터베이스는 부족하고, 반면에 정보자원은 그 비율에 있어서 많은 편에 속하지만, 전체적인 측면에서 보면 정보서비스나 정보자원이 매우 부족한 분야이다. 특히 첨단과학분야에는 공공데이터베이스뿐만 아니라 정보자원도 없는 실정이다.

2) 정보개발 현황의 분석

본 연구에서 실시한 국내 데이터베이스, 국가정보자원, 그리고 해외 데이터베이스에 대한 조사는 전술한 바 있듯이 제한적으로 이루어져 있다. 즉 국내 데이터베이스는 공공유통망을 통하여 일반인에게 공개되어 있는 모든 데이터베이스를, 즉 그 정보원이 국가 행정기관 또는 공공기관에 속하는지에 대한 여부는 묻지않고, 망라하는 공공 데이터베이스에 대한 한국데이터베이스진흥센터의 조사에 근거를 둔 반면에, 국가정보자원은 국가기관 및 공공기관이 소유하고 있는 자료중 공공정보로써 공개가능하고 개발가능한 것에 제한하고 있음으로 해서 그들간의 비교가 일반화되기는 힘들다. 또한 해외 데이터베이스의 현황에 대한 조사도 Gale Research Inc.에서 수집하여 발간한 Gale Database Directory[1995]와 The Federal Data Base Finder[1995]에 근거를 두고 이루어지는 제한이 있어 국내 데이터베이스 현황 그리고 정보자원과의 비교가 일반화되기는 힘들다는 판단이다. 그러나 본 연구의 이러한 제한점을 인식하면서, 본연구의 목적인 공공정보개발 사업을 위

한 정보개발현황을 파악하는 것은 부분적으로 가능케한다.

정보개발 현황에 대한 분석을 위하여 여기서 정보개발의 취약분야, 풍요분야, 그리고 중점분야의 개념을 사용한다. 먼저 취약분야란 기존 국내 정보개발이 해외에 비해서 취약한 분야이고, 풍요분야란 개발된 정보서비스에 비해서 상대적으로 정보자원이 풍부한 분야이다. 중점분야는 앞의 두분야를 고려하여 기존의 정보개발이 취약하지만 개발가능한 자원이 풍요한 분야를 말한다. 이러한 기준들은 정보의 주제별 분야에 있어서 먼저 대분류와 중분류의 측면에서 적용된다. 다음으로는 보다 구체적으로는 각 정보자원에 대한 개발의 우선순위는 이를 정확하게 그 순서를 나열하기 보다는 각 자료에서 그 우선순위를 파악할 수 있는 요소를 중심으로 개괄적으로 제시함으로써 판단의 기초를 제공하고자 한다.

(1) 대분류별 분석 - 정보개발의 대분류별로 분석해보면, 우선 국내 공공 데이터베이스가 해외 데이터베이스에 비해서 상대적으로 취약한 부분은 과학/기술, 정보기술/통신, 사회일반, 산업, 경제/경영, 증권/금융, 생활, 교육/도서, 문화/레포츠 분야의 순으로 분석되어진다. 다음으로 국내에서 개발되어진 공공 데이터베이스에 비해서 상대적으로 정보자원이 풍부한 분야를 보면, 사회일반, 산업, 과학/기술, 경제/경영, 문화/레포츠, 생활, 교육/도서, 정보기술/통신, 증권/금융 분야의 순으로 분석되어진다.

따라서 개발의 취약한 분야와 자원의 풍요한 분야의 분석을 통한 각 분야에 가중치를 일정한 간격척도로 배정하여 측정할 때, 차후 정보개발의 중점분야로 제시될 수 있는 것은 앞의 표에서와 같이 과학/기술, 사회일반, 산업, 경제/경영, 정보

〈표 17〉 정보개발의 구분별 현황분석

구분	순 위
개발 취약 분야	과학/기술, 정보기술/통신, 사회일반, 산업, 경제/경영, 증권/금융, 생활, 교육/도서, 문화/레포츠 분야 등의 순
자원 풍요 분야	사회일반, 산업, 과학/기술, 경제/경영, 문화/레포츠, 생활, 교육/도서, 정보기술/통신, 증권/금융 분야 등의 순
차후 중점 분야	사회일반, 과학/기술, 산업, 경제/경영, 정보기술/통신, 생활, 문화/레포츠, 증권/금융, 교육/도서 분야 등의 순

기술/통신 분야 등의 순이다. 특히 여기서 정보기술/통신 분야의 경우는 현재의 개발된 정보서비스뿐만 아니라 정부에서 제공하는 자원조차도 미비한 것으로 나타나고 있다. 이러한 것은 앞으로의 국제화, 세계화, 그리고 정보화를 추구하는 정부의 국정정책에 부합하기 위해서도 중점적으로 개발 육성되어야 할 분야로 판단된다.

(2) 중분류별 분석 - 정보개발의 중분류별로 분석해보면, 먼저 국내 데이터베이스가 해외 데이터베이스에 비해서 상대적으로 취약한 분야를 보면, 환경/공해, 투자신탁, 법률, 자원/에너지, 노동, 과학/기술일반, 정보통신, 교통정보, 경제/경영일반 등의 순으로 나타난다. 또한 국내 데이터베이스의 수에 비해서 상대적으로 정보자원이 풍부한 분야는 자원/에너지, 노동, 환경/공해, 사회복지, 유통, 지역, 교통정보, 농림축산업, 통일/국제, 정치/행정 등의 순으로 나타난다.

따라서 개발의 취약한 분야와 자원의 풍요한 분야의 분석을 통한 각 분야에 가중치를 일정한 간격척도로 배정하여 측정할 때, 환경/공해, 자원/에너지, 노동, 교통정보, 과학/기술일반, 정치/행정, 법률, 사회복지, 유통, 지역, 경제/경영일반, 정보통신, 전기/전자, 특허/상표 등의 순으로 판

단된다.

(3) 정보자원별 분석 - 차후 정보개발 사업을 위한 정보자원별 우선분야에 대해서는 고려되어야 할 사항들이 많고, 또 보는 관점에 따라서 다양할 수 있다. 본 연구에서는 각 정보자원의 갱신주기, 주요 이용자, 표현형태, 그리고 정보개발의 시급성을 주요 판단기준으로 제시한다. 즉 갱신주기에 의해서 자료변화의 신속 또는 긴급성을 판단할 수 있으며, 주요 이용자에 의해서 정보개발에 따른 혜택의 범위를 판단할 수 있으며, 자료의 표현형태에 따라서 차후 데이터베이스의 표현형태가 가름된다(예, 멀티미디어). 정보개발의 시급성은 정보자원을 보유하고 있는 정부기관에서 판단한 것으로 우선분야를 결정하는데 도움이 될 것이다.

Ⅶ. 국가 공공데이터베이스의 개발 과제

정보개발의 현황분석을 토대로 차후 국가의 공공데이터베이스의 개발 과제는 다음에서와 같이 기초과제, 중점과제, 그리고 개척과제로 구분하여 제시될 수 있다.

먼저 기초과제란 해외에서는 해당분야 정보개발의 중요성이 인식되어 많은 정보서비스가 이루어지고 있으나 국내에서는 현재 정보서비스가 이루어지지 않은 분야이다. 여기에 해당되는 대표적인 분야는 첨단과학분야이다. 첨단과학분야는 본 조사에서 보면, 해외에는 61개의 데이터베이스를 볼 수 있는데 반하여, 국내에는 이에 관한 데이터베이스도 없으며, 사실 자료들도 공개되지 않고 있는 실정이다.

둘째로 중점과제는 해당분야의 정보서비스가 이루어지고 있지만, 그 규모가 해외에 비하여 취약한 반면에 개발가능한 정보자원이 많은 분야에 우선순위를두어 중점적으로 개발을 유도하는 과제이다. 이를 위해 우선적으로 정보의 대분류를 따라 중점분야를 선정하고, 각 분야의 중분류별로 개발의 우선과제를 선정하는 것이 제안될 수 있다. 본 연구에서 중점과제를 앞의 정보개발현황분석결과에 의하여 표 4에서 제시하고 있다.

마지막으로 개척과제는 본 연구의 조사범위에 속하지 않는 과제이다. 전술하였듯이 본 연구의 조사는 정부기관 및 공공기관에서 이미 자료를 수집하여 이를 대중에 공개하고 있는 정보자원을 중심으로 하고 있다. 그러나 최근 자료의 중요성이 인식되고 공공정보로써 개발되어야 할 필요성과 시급성이 높게 인정되고 있는 분야들이다. 이러한 정보에 대한 개발 또한 중요한 부분으로 다루어져야 한다는 것이다. 대표적인 實例로 島嶼에 대한 정보를 들 수 있다. 1994년 기준 486개의 유인도를 포함하여 약 3,200개의 도서지역을 국가경쟁력을 갖는 지역으로 만들기 위한 도서개발의 방향정립과 전략수립에 요구되는 자료로, 각 도서의 입

지적 여건(인근 도서와의 관계, 해안 중심지역과의 관계, 자연생태적 조건), 인구구조, 면적규모, 교통여건, 산업구조 등에 대한 자료가 현재 수집되지 못한 상태이지만, 이러한 자료들은 수집되어 공공정보로서 개척되어야 할 중요한 분야인 것이다.

Ⅶ. 결 론

본 연구는 총괄적 정보개발계획을 수립하기 위하여 국가의 정보개발 현황을 분석할 수 있는 일명 '정보개발지도'를 제작하여 국가 공공데이터베이스 개발현황을 분석하고 개발과제를 파악하고자 하였다. 이는 현재 국내에 개발되어 서비스를 제공중인 정보, 사업의 지원자금으로 개발을 추진 중에 있는 정보, 혹은 아직 미개발되었지만 공공정보로써 개발이 가능한 공공기관의 정보자원을 조사하여 세부 분야별로 분류하여 하나의 도면에 나타내고, 또 실질적으로 우리의 공공정보의 개발, 서비스의 역사는 짧기 때문에 선진화된 외국의 정보개발서비스들을 지도에 나타냄으로써 향후 공공정보개발의 방향을 파악할 수 있게 한다. 즉 이 지도는 향후 공공데이터베이스 개발과제의 파악을 위한 기반자료로 활용될 수 있으며, 정보개발 지도에 따라 년차별 개발분야 및 개발우선순위 등이 결정될 수 있을 것이다. 구체적으로 본 연구는 정보개발지도에 대한 체계적 연구를 비롯하여 국내 공공정보개발의 현황을 파악하기 위한 국내 및 해외의 데이터베이스현황조사 및 국내 정보자원 조사와 향후 공공정보개발사업을 위한 건의 등을 포함하였다.

본 연구는 정보자원 및 데이터베이스 현황을 조

〈표 18〉 공공데이터베이스의 개발 과제

구분	중점과제	
기초과제	첨단과학분야	
중점과제	과학/기술분야	환경/공해, 과학/기술일반 분야
	사회일반분야	사회복지, 법률, 정치/행정 분야
	산업분야	자원/에너지, 유통 분야
	경제/경영분야	노동, 경제/경영일반
	정보기술/통신분야	정보통신, 정보기술 분야
개척과제	국토개발분야 (예, 도서개발 정보)	

사하는 가운데 많은 한계점을 갖고 있다. 가령 정보자원조사의 경우, 그 조사의 대상이 정부기관 및 공공기관에 한정됨으로써 광의의 의미로써 공공데이터베이스 개발을 위한 정보자원에 대한 망라적 조사에는 이르지 못하였다. 또한 공공기관에 대한 자원조사도 관련 공공기관의 비협조로 인하여 충분한 정보자원이 수집되었는지에 대한 의구심도 없지 않다. 해외 데이터베이스의 경우에도 그 조사가 단기간에 수행되기 위해서 기존의 조사 자료집을 이용함으로써 그 한계점을 가져왔다고 할 수 있다. 또 하나의 한계점은 본 연구가 현실 지향적인 연구에 그친 것이다. 즉 현존하는 자료에 대한 조사에 치중되었고, 개발가능한 자료에 대한 접근을 할 수 없었다. 따라서 이러한 정보자

원에 대해서는 지속적인 개척이 요구되며, 지속적인 국가정보화 사업을 위해서도 가용 정보자원에 대한 지속적인 조사가 요구된다. 이를 위해서는 공공정보 또는 공공데이터베이스에 대한 분류법에 대한 체계적 접근이 연구되어야 한다. 연구 목적에 맞는 분류법을 제시하기 위해서는 그 분류가 제작자 위주의 분류가 고려될 필요가 있음에도 불구하고, 기존의 연구가 부족하고 연구결과의 단기적 안목에서의 활용을 위해서 사용자 위주의 분류법을 사용함으로써 본 연구에서의 분류법은 한시적 효과만을 가져올 것이 예상된다. 차후의 연구에서는 우선적으로 공공데이터베이스에 대한 체계적인 개념적 접근과 분류법이 확립되어야 할 것이다.

〈참고문헌〉

- Davenport, Thomas *Process Innovation: Reengineering Work through Information Technology*, Harvard Business School Press, Boston, 1993.
- Elmasri, Ramez and Shamkant B. Navathe. *Fundamentals of Database Systems*, The Benjamin/Cummings Publishing Company, Inc. Redwood City, CA. 1994.
- Hammer, Michael and James Champy, *Reengineering the Corporation: A Manifesto for Business Revolution*, HarperCollins Publishers, Inc., New York, 1994.
- Hernon, P. "Government Information", in C. R. McClure et al, *United States Government Information Policies : Views and Perspectives*. Ablex Publishing Corporation, Norwood, NJ., 1989.
- Lesko, Matthew (ed.), *The Federal Database Finder*, Information USA, Inc., Kensington, MD. 1995.
- Nolan, Kathleen Lopez (ed.), *Gale Directory of Database, Volume 1: Online Databases*, Gale Research Inc., Detroit, MI., 1995.
- Random House, *The Random House Dictionary of the English Language*, 2nd ed., Random House, New York, 1987.

- 고영만 외, 「데이터베이스 표준분류 및 정보검색 지침서 표준안을 위한 기초연구(DPC 표준화 94-03)」, 1994.
동아출판사, 「동아대백과사전」, 동아출판사, 서울, 1987.
- 이국희, “공공데이터베이스 가치의 측정,” 「데이터베이스 저널」, 제2권 1호 1995, pp. 67~89.
- 이재홍, “데이터베이스 정책의 향후 추진방안,” 「제3회 서울 국제 데이터베이스 컨퍼런스: I. DB Policy & Issues」, 1995. pp. 51~63.
- 정보통신부, 「공공데이터베이스 개발·보급사업 개선방안」, 1995. 8.
- 정부간행물제작소, 「정부간행물목록」, 1995.
- 정운, “온라인 데이터베이스 산업의 이슈,” 「제3회 서울 국제 데이터베이스 컨퍼런스: I. DB Policy & Issues」, 1995. 9.
- 한국데이터베이스진흥센터(DPC), 「'95년도 공공데이터베이스 개발 과제 조사」, 1994. 10.
- 한국데이터베이스진흥센터(DPC), 「알기쉬운 한국 데이터베이스 목록 1995」, 1995. 3.
- 한국데이터베이스진흥센터, “국내의 데이터베이스산업의 동향과 전망(Ⅱ),” 「데이터베이스월드」, 1995. 7a, pp. 41~53.
- 한국데이터베이스진흥센터(DPC), 「정보개발지도구축계획(안)」, 1995. 7b.
- 한국데이터베이스진흥센터(DPC), 「정보개발지도 제작 연구」, 1995. 10.
- 한국통신, 「'95공공 데이터베이스 개선(안)」, 1995. 7.