

공공데이터베이스개발을 위한 제언 — 체신부 공공데이터베이스개발사업을 보고 —

(Recommendation for Developing Public Database)



임 성택
고려대학교 경영정보학과 조교수
Rim, Seong-Taek Ph. D.
Assitant Professor, Korea University

목 차

1. 공공데이터베이스란 무엇인가?
2. 체신부 (한국통신) 공공데이터베이스개발사업
 - 2.1 사업개요
 - 2.2 추진체계및 절차
 - 2.3 개발과제 평가및 방법
 - 2.4 과제선정절차 및 방법
 - 2.5 개발업체의 선정
3. 공공데이터베이스개발사업의 문제점및 개선방향
4. 맺음말

공공데이터베이스개발을 위한 제언

데이터베이스산업은 정보통신분야의 핵심분야이며 정보화의 기반산업으로서 그 중요성이 날로 증대되고 있으며, 그 자체로도 고부가가치를 지닌 유망산업으로서 부존자원이 부족한 우리로서는 국가적인 차원에서 전략적으로 육성해 나가야

할 분야중의 하나이다. 뒤늦게 출발한 우리의 데이터베이스산업이 국제무대에서 경쟁력을 갖고, 이들과의 격차를 줄이기 위해서는 이미 이들 선진국에서 경험했던 시행착오(실패)를 겪지않고 보다 효과적으로 데이터베이스산업을 육성해나가는 것이 중요하다고 하겠다. 데이터베이스후발국으로서 어떻게 하면 우리가 가장 효과적으로 데이터베이스산업을 발전시킬 수 있을까? 이에 대한 해결방안중의 하나가 바로 정부주도하의 공공데이터베이스를 개발하는 것이라고 하겠다.

때마침 정부에서도 공공데이터베이스개발에 정부지원의 필요성을 인식하고 공공데이터베이스개발사업을 발표하게 되었다. 그 주내용은 800억 원의 예산을 투입하여 4년간 매년 100정도의 공공 데이터베이스를 개발하여 공중망에 접속하여 일반에 개방한다는 것이다. 개발과제는 국가정책적인 차원을 고려, 한국데이터베이스진흥센터에서 국내데이터베이스개발 수요조사를 실시하고 이를 근거로 데이터베이스개발과제를 선정하며 체신부에서 선정된 과제를 확정하고 한국통신은 개발보급을 맡는다는 것이다. 이에따라 한국통신은 4년간 공공데이터베이스를 개발보급하는 틀이라할 수 있는 공공 데이터베이스계획을 확정했고 지난 5월 금년에 개발할 과제를 확정했다. 이러한 데이터베이스산업에 대한 정부의 관심은 뒤떨어진 우리의 데이터베이스산업을 고려할 때 아주 바람직한 현상이라고 할 수 있다.

이글의 목적은 현재 추진중인 체신부 공공데이터베이스개발사업을 바탕으로 효과적인 공공데이터베이스개발 방안을 제시하는 데 있다. 이는 다음 네가지 질문에 의해 구체화 되었다.

- ▷ 공공데이터베이스란 무엇인가?
- ▷ 왜 정부주도의 공공데이터베이스개발사업이 필요한가?
- ▷ 정부가 하여야할 역할은 무엇인가? 그렇다면 민간사업자는?

▷ 과연 정부는 현재 추진중인 사업에서 이러한 역할을 충분히 수행하고 있는가?

1. 공공데이터베이스란 무엇인가?

최근들어 부쩍 ‘공공데이터베이스’란 용어가 많이 사용되기는 하지만 이에 대한 정확한 정의 없이 사용하는 이마다 각기 다른 의미로 사용하고 있는 실정이다. 따라서 공공데이터베이스개발방안에 대해 살펴보기 전에 공공데이터베이스의 개념을 정리해 볼 필요가 있다.

공공데이터베이스에 대한 정의는 조동성이사의 글과 체신부 사업계획서에서 찾아 볼 수 있다. 한화 경영기획실의 조동성이사는 ‘공익데이터베이스’라 하여 ‘정보의 사회적 유의성(social significance)’에 의해 정의했다. 정보의 사회적 유의성(social significance)이란 사회구성원(개인, 조직)이 그들의 의사를 결정하며 그것을 적용하여 새로운 인식을 형성하고 발전을 도모할 수 있는 데 매우 요긴한 것으로 정의했다. 이러한 ‘사회적 유의성’이란 기준하에서 그는 공익데이터베이스를 1) 가능한 한 다수에게 혜택이 돌아갈 수 있는 데이터베이스, 2) 독점되었을 경우 사회구성원에게 피해가 돌아올 데이터베이스, 3) 요구는 많으나 수익성이 적어서 상품화의 가능성이 적은 데이터베이스, 4) 정보통신서비스에 친숙함을 높일 수 있는 데이터베이스, 5) 데이터베이스구축에 소요되는 시간과 자금면에서 실현 가능한 데이터베이스등으로 현실화(operationalize)하였다. 그는 또한 이러한 기준에 따라 공익데이터베이스의 유형을 행정정보, 의료정보, 안전정보, 산업정보, 교육정보, 문화정보, 생활정보등으로 구분하였다 (조동성, 1993).

〈공익데이터베이스의 유형〉

행정정보 : 주민정보, 세무행정정보, 정책제도 정보, 통계정보, 병무정보

- 의료정보 : 구급의료정보, 예방접종, 주택진료, 복지정보
- 안전정보 : 재해방지정보, 소비자지원정보, 환경오염정보, 교통관제
- 산업정보 : 첨단기술/특허정보, 기술개발지원정보
- 교육정보 : 도서관 온라인시스템, CAI서비스, 시민대학강좌정보, 진학
- 문화정보 : 문화행사정보, 지역문화재정보, 마을정보
- 생활정보 : 상품정보, 물가정보, 기상정보, 취업구인, 가정의례

체신부 공공데이터베이스개발사업계획서에서는 공공데이터베이스를 개념적으로 경제사회적으로 필요하나 민간이 독자적으로 개발하기 곤란한 데이터베이스, 다수의 국민이 필요한 정보(공공부문, 민간부문 포함)를 수록한 데이터베이스, 시간과 교통등 사회적 물류비용을 줄일 수 있는 데이터베이스로 정의하고, 구체적인 예로서, 산업경쟁력강화를 위한 산업기술등 전문데이터베이스, 국민의 편익증진을 위한 생활데이터베이스, 지역간 정보격차 해소를 위한 지역데이터베이스, 장애인 복지증진을 위한 데이터베이스등을 들고 있다.

공공데이터베이스에 대한 앞의 두가지 정의는 데이터베이스가 사회구성원에게 주는 혜택(benefit), 즉 데이터베이스의 공익성에 의해 정의했다. 명확한 공공데이터베이스에 대한 개념정의는 반드시 필요하지만 이는 다음으로 미루기로 하고 이 글에서는 일단 정보의 공개성과 사회적 유익

성을 강조하여 “공공데이터베이스란 자료의 공개가 가능하고, 사회의 구성원(조직, 개인)에게 유익함을 줄 수 있는 데이터베이스”라 정의하기로 하자. 지금부터는 과연 이러한 공공데이터베이스가 정부주도하에 이루어지는 것이 바람직한지에 대해서 살펴보기로 하자.

2. 체신부 공공데이터베이스개발사업

2.1 사업개요

- 데이터베이스산업육성및 발전기반 구축
 - 민간 데이터베이스사업자에 대한 시장조성
 - 공공부문의 정보공개촉진
 - 산업경쟁력강화를 위한 산업데이터베이스 구축지원
 - 지역간의 정보격차 해소를 위한 지역데이터베이스 구축지원
- 공공데이터베이스 개발규모

2.2 추진체계및 절차

한국데이터베이스진흥센터(DPC)는 정부및 공공기관의 보유자료 실태를 파악, 매년도 개발대상 공공데이터베이스의 종류및 범위를 정하여 정부에 제출하며, 이를 위해 한국데이터베이스진흥센터내에 산학연의 전문가로 구성된 공공데이터베이스개발추진협의회를 구성운영하고 있으며 체신부는 매년도 개발할 공공 데이터베이스를 확정하여 한국통신에 통보하며 한국통신은 개발과제를 공고하여 데이터베이스개발사업자를 선정하

구 분	'94	'95	'96	'97	계
공공데이터베이스개발수(개)	100	100	100	100	400
소요자원(억원)	200	200	200	200	800

* 개발과제수는 신속적으로 운영할 계획임

여 공공데이터베이스 개발보급을 추진하고 있다.

2.3 개발과제 평가 및 선정

과제 선정 기준

- 전체과제 선정기준
 - 국가 데이터베이스산업의 정책적 필요성
 - 사회적/산업적 시급성
- 개별과제 선정기준
 - 데이터베이스산업기술적 강화
 - 투자환경 조성
 - 데이터베이스 수요 확대
 - 개발효과성
- 필수 충족기준
 - 기존 데이터베이스와의 중복 여부(유통 및 개발 중인 데이터베이스)
 - 공개성(개발 완료 데이터베이스의 공개 및 정보통신망 유통)
 - 사업목적과의 적합성(불건전정보 등)

과제 평가 항목

- 평가항목 선정 절차
 - 과제 선정 기준, 평가범위, 평가범주를 결합하여 평가기준을 도출
- 평가범위 및 기준
 - 정보 : 정보 가치
 - 정보원 : 신뢰성
 - 개발 : 최소한의 품질을 가진 데이터베이스 개발 능력
 - 운영 : 신뢰성
 - 이용자 : 이용자 만족도
- 평가범주
 - 과제 개발의 타당성 : 수요 측면에서 정보의 이용 가치 및 효익, 긴급성, 공공성, 시장성 등을 기준

- 실현가능성 : 공급 측면에서 정보의 안정적 수집, 지속적인 갱신, 가공처리 능력, 개발 완료 데이터베이스의 운영 및 유지 능력 등을 기준으로 개발 실현 가능성

- 평가 기준 설정
 - 조사양식서에 의거하여 평가항목을 구성하여 조사
(통계적 분석 방법과 비통계적 평가항목 조사 방법)
- 평가항목
 - 평가항목의 구성에
 - 원시자료의 지속적 제공 가능성 여부
 - 정보수집체계의 안정성 여부 등

2.4 과제 선정 절차 및 방법

- 1차 탈락과제 선별
 - 필수 충족기준에 부합하는 과제를 선정
 - 기타 과제는 개발대상에서 일단 제외
- 2차 평점제 및 비통계적 평가
 - 평점제 : 과제의 타당성과 실현가능성을 각각 통계적/비통계적 방법으로 측정하여 그 우선순위를 과제 평점으로 도출
 - 비통계적 평가
 - 과제의 특성을 반영기 위해 비통계적 평가항목을 추가
 - 평가항목에 기초한 통계적인 평점제의 적용한계를 보완
- 3차 비율조정
 - 정책적으로 육성이 필요한 분야
 - 시장/기술의 파급효과가 큰 분야
 - 4년간의 단계별 투자의 전체적 효과를 최대화
 - 전문가용과 일반범용 데이터베이스
 - 매년 중점지원분야를 선정하여 사업의 효과성 향상
 - 선정된 유사과제를 기관별로 특화하여 조정 통합 추진

○ 과제 최종심 및 탈락과제 재심의 조정

- 전체 제안과제에 대한 일괄검토/조정
- 타당성과 실현가능성에 의한 적정과제 도출

2.5 개발업체의 선정

'93년 12월 소비자보호원, 특허청 등 114개 기관에서 제안한 512개의 과제를 대상으로 앞에서 설명한 과제 선정기준에 의하여 93개의 과제가 최종적으로 확정했다. 개발과제의 분야별 비율은 일상생활이 15%, 경제산업이 45%, 학술교육이 30%, 기타가 10%의 순이다.

입찰참가자격

참신한 기술을 가진 사람이 많이 참여할 수 있도록 참가자격을 대폭 완화하여 다음의 참가자격을 갖추면 누구나 참여할 수 있게 했다.

- 전기통신사업법 시행규칙 제3조에 의한 부가통신사업자
- 소프트웨어개발촉진법 제5조의 소프트웨어 개발사업자
- 국가기관, 정부투자기관, 언론기관, 협회, 학회 등

선정방법

개발업체의 선정은 제안서에 입각한 평가에서 선정기준 이상의 점수를 획득한 업체중 최고득점자로 한다. 이 제안서의 평가는 투자효과를 극대화하기 위한 조건들, 예를들면, 원시자료확보의 계속성 보증, 정보현행화 보증여부, 개발된 데이터베이스의 안정적 운용방안등과 같은 필수조건을 충족하는 제안서에 한해 평가요소별 득점에 의하고 여기에 추가로 개발비가 한국통신이 정한 범위내에 있어야 한다.

3. 공공데이터베이스개발사업의 문제점 및 개선

방향

앞에서 우리는 우리가 데이터베이스산업분야의 후발국이란 점에서 정부주도하에 공공데이터베이스개발이 우리의 데이터베이스산업발전을 위해 아주 바람직하다고 했다. 하지만 현재 체신부의 공공데이터베이스개발사업이 이러한 장점을 충분히 살리기 위해서는 다음 몇가지 문제점을 개선해야 할 필요가 있다.

첫째, 외국의 데이터베이스산업에 대한 철저한 분석이 이루어지지 않았다는 들 수 있다. 정부주도 공공데이터베이스개발의 장점중의 하나로 선진국의 과오를 되풀이 하지 않아 발전속도를 빨리할 수 있다는 것이었다. 그러기 위해서는 선진 외국의 대표적인 데이터베이스들 및 기술동향등을 면밀히 분석하는 것이 필요하다. 이는 단순히 데이터베이스, 공급업자, 개발업자의 숫자등과 같은 통계적인 자료만을 의미하는 것이 아니고, 각 분야별로 성공적인 데이터베이스사례(아니면 실패사례)에 대한 철저한 분석이 있어야 한다. 다시 말하자면 대표적인 성공데이터베이스들의 중요성공요인들(Critical Success Factors)를 면밀히 파악하여 이들을 우리에게도 적용할 수 있는 지를 살펴보는 것이 좋을 것 같다. 과연 이들 데이터베이스들을 성공하게 만든 사회적, 기술적 기반은 무엇인가? 이러한 데이터베이스를 개발할 때 예상되는 어려운 점은 무엇인가? 어떠한 데이터베이스개발, 유통체제를 갖고 있는가? 이런 철저한 분석이 선행된 후에 과제를 선정할 경우 그 과제를 수행할 수 있는 민간사업자도 선정/관리하기도 쉬울 뿐더러 아주 쓸모있는 데이터베이스를 만들 수 있을 것이다.

둘째, 개발과제가 너무 많아 잘못하면 수준이하의 데이터베이스를 양성하고, 군소업체들을 난립을 초래하여 데이터베이스산업을 활성화하겠다는 본래의 목표를 흐리게 할 수 있다는 것이다. 원래 과제의 수를 다양하게 한 것은 민간업자들

로 하여금 데이터베이스개발의 붐을 조성한다는 취지로 생각되지만, 이러한 소비자들의 데이터베이스수요창출과 민간업자들의 데이터베이스개발의욕은 몇개의 시범적인 데이터베이스가 성공리에 개발 이용됨으로서 만들어질 수 있으리라 생각한다. 이를 개선하기 위해서는 현재 개발과제의 큰 분류기준은 그대로 유지하되, 개발데이터베이스의 과제수를 대폭 줄였으면 한다. 예를 들면 국가적으로 아주 시급하지만 민간업자가 참여하기 어려운 데이터베이스 한두개와 상업적으로 성공이 가능한 분야(외국의 사례분석에서 나타난) 몇개의 pilot 데이터베이스에 집중투자하는 것도 고려해볼 필요가 있다. 이러한 pilot 데이터베이스가 상업적으로 크게 성공한다면 데이터베이스개발붐은 자연적으로 조성되리라 생산한다.

셋째, 현재의 문자, 온라인 위주의 데이터베이스개발보다는 멀티미디어, CD-ROM 등 각종 첨단정보기술을 활용할 수 있는 데이터베이스구축을 유도 했으면 한다. 지난 20년동안 미국의 생활정보시장을 지배해오던 CompuServe가 시장 점유율면에서 Prodigy에 의해 2위로 밀려났다. Prodigy가 등장과 함께 1위의 자리를 확보할 수 있었던 것은 CompuServe와는 달리 그래픽 인터페이스를 바탕으로 사용자기호에 맞는 서비스를 제공한다는 점이다. 이는 최근 급부상한 America Online도 마찬가지이다. 이러한 그래픽의 중요성에도 불구하고 CompuServe를 비롯한 기존의 서비스들이 이를 따라갈 수 없었던 것은 이미 오래 전의 기술로 만들어진 기초적인 기능들을 근본적으로 바꾸는 데는 큰 투자와 위험이 따르기 때문이다. 또한 최근 CD-ROM 데이터베이스의 발달은 데이터베이스의 사용을 전문가집단에서 일반에게로 대중화하는 데 기여했다. 우리는 기왕이면 먼 장래를 내다보고 이러한 첨단 정보기술을 이용했으면 한다.

넷째, 위에 것과 연관되는 것이지만 데이터베이스는 그것을 유통시키는 통신망과 불가분의 관

계에 있다. 다시말해 데이터베이스는 정보서비스 상품이며 동시에 통신서비스 상품이다. 따라서 현재 초고속정보통신망은 21세기 고도정보사회를 주도할 국가기반구조로서 새로운 사회간접자본으로 부상하고 있는 데, 이러한 초고속, 대용량의 멀티미디어 정보수요에 대비하는 데이터베이스개발이 필요하다.

이외에도 구체적인 사업추진계획을 살펴보면, 과제도출을 위한 수요조사를 실시할 경우 관심있는 사람들이 모두 참여할 수 있도록 충분한 홍보 및 시간이 필요하다든지, 여러 이해관계당사자중 공급자(supplier)의 입장이 주로 반영되고 고객(customer)의 목소리가 도외시되고 있다든지, 사업관련 업무분담이 체신부, 한국통신, 한국데이터베이스진흥센터등으로 분산되어 있어 업무의 일관된 정책이 어려웠다는 점등을 들 수가 있다.

4. 맺음말

결론적으로 말하자면 우리가 데이터베이스산업의 후발국이란 점에서 정부주도하에 공공데이터베이스를 개발하는 것은 아주 바람직한 일이라 할 수 있다. 이러한 의미에서 체신부(한국통신)의 공공데이터베이스개발사업은 아주 환영할만한 일이라고 하겠다. 하지만 이 사업이 충분한 성과를 지니기 위해서는 정부는 이 공공데이터베이스 개발사업에서 정부가 하여야할 역할(리더십과 비전을 제시, 이해조정, 기술적인 기반구조 확립)을 충실이 해야할 것이다.

또한가지 언급하고 싶은 것은 데이터베이스도 하나의 상품이므로 궁극적으로는 민간기업의 참여에 의한 자유시장원리에 의해 개발되어야 한다는 것이다. 따라서 지금이라도 공공데이터베이스 개발은 정부가 주도하되(공공기관의 정책사업에 의해), 과제의 선정이나 데이터베이스개발은 철저히 상업성과 경쟁력을 바탕으로 이루어져야 한다는 것이다.