

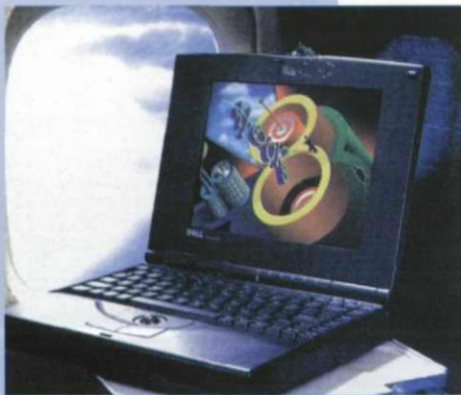
데이터베이스 정책개발을 서둘러야

최동휘

한국통신 S/W연구소장

미래라는 말과 정보화라는 말은 뭘래야 뭘수 없는 말이 되었다. 그리고 정보화의 원동력은 잘 축적되고 관리되어 있는 데이터베이스로부터 나온다는 사실에 이의를 제기하는 사람도 없다. 이렇게 미래 사회의 산실이라고 할 수 있는 데이터베이스란 무엇이고 그것이 산업적 측면에서 어떻게 존재하는가? 보통 데이터베이스란 컴퓨터 등 정보처리 장치를 써서 문자, 기호, 도형, 음성, 화상, 영상 등으로 구성된 자료를 체계적으로 수집 축적하여 이용자가 편리하게 이용할 수 있도록 한 자료의 집합체를 말한다.

여기서 데이터베이스를 만드는 일, 즉 자료를 수집, 처리, 가공하여 정보를 제작, 생산한다는 것은 산만하게 흩어져 각양 각색으로 우리 생활 속에 파고 들어와 있는 각종의 원시자료를 수집하고, 체계적으로 정리하여 전산매체로 가공함으로써 이용자들이 요구하는 정보를 제공할 수 있도록 준비하는 작업 일반을 말한다. 이렇게 생산되어진 "전산매체로 된 자료" 즉 데이터베이스를 데이터베이스 관리시스템(DBMS)에 알맞게 구현하여 이용자의 요구에 부응할 수 있는 서비스 환경을 통해 유통시키는 것을 데이터베이스 서비스 제공 즉, 데이터베이스 유통이라고 말한다. 여기서 우리는 데이터베이스 산업의 특성을 읽어볼 수 있다. 먼저 데이터 즉 자료란 문자, 기호, 음성, 화상, 영상 등 우리가 보고 느낄 수 있는 모든 수단으로 이루어진 복합체이다. 어떤 자료



는 문자를 중시하는 조합으로, 어떤 자료는 음성을 중시하는 조합으로, 어떤 자료는 화상을 중시하는 조합으로 이루어지고 있다. 그리고 이용자는 자기의 기호에 따라 자기가 빨리 알아볼 수 있는 자료를 원하고 있다.

다음으로 다양한 정보가 우리의 일상 생활에 나도 모르는 사이에 각양 각색으로 파고 들어와 있다는 사실이다. 마치 공기나 물처럼. 따라서 정보는 대가없이 제공되는 것처럼 인식되는 경향이 있는 반면 다른 한편으로는 “사람들이 필요로 하는 것은 보편적으로 제공되는 정보가 아니라 자기가 필요로 하는 특수한 정보”인 경우가 많다. 여기서 정보는 공기나 물처럼 무료로 제공되는 것이고 필요한 것만 돈을 주고 사게 되는 상업성 정보의 다품종 소량수요 문제가 제기된다. 마지막으로 똑같은 정보라도 이용자에 따라서 서로 다른 효용가치를 제공하고 있는 특징이 있다. 정보처리 장치의 이용(예 PC 등)에 대한 숙련도, 어떤 분야에 대한 이용자의 전문성 등에 따라 이용자가 썬하는 정보의 가치는 서로 다르게 나타난다는 점도 정보 산업의 특성에 영향을 미치게 된다.

이러한 특징을 갖고 있는 데이터베이스를 생산이라는 측면에서 검토하고자 하는 것이 본고의 주제이다. 이러한 관점에서 우리는 데이터베이스의 산업적 위치, 데이터베이스의 생산구조, 데이터베이스의 생산 능력 및 데이터베이스가 상품이라는 측면에서 보는 유통형식 등을 검토할 수 있을 것이다.

첫째로 데이터베이스 산업은 어떻게 존재하는가의 문제다. 데이터베이스는 프로그래밍, 인터페이스, 코딩, 문자처리, 언어처리 등을 중심으로 하는 소프트웨어 산업의 가장 폭넓은 응용형태로 나타난다. 전자가 기술이라면 후자는 이것을 응용하는 사업의 형태로 나타나며, 전자는 “생산단지, 유통단지, 유통구조 개선 등 제도중심의 지원

형태로” 주도 그룹이 사업을 부추기는데 반해 후자는 “생산, 유통, 보급 등에 자금을 직접 지원”하는 형태로 사업을 부추긴다. 아울러 전자는 원천기술이라는 차원에서 업체 중심으로 스스로 참여하는 사업의 성격이 강한데 비하여, 후자는 사업을 주도하는 그룹이 개발된 기술을 활용하여 특정 목적을 수행하기 위한 사업으로써의 성격이 강하다.

예를 들어 한글을 정보처리하기 위하여 한글을 표준화하고 입출력 기기를 생산하는 것이 전자의 예라면, 정보화 사회 구현을 위하여 정부가 공공 데이터베이스 사업을 일으킨다든지, 특정 재벌그룹이 자사의 경영혁신을 위하여 BPR 형태의 사업을 일으키는 것은 후자의 예이다. 여기서 우리는 소프트웨어 산업의 지원이 데이터베이스 산업의 지원이나? 데이터베이스 산업의 지원이 소프트웨어 산업의 지원이나? 데이터베이스 산업의 지원이 소프트웨어 산업의 지원에 어느 정도의 영향을 미치는가? 라는 명제에 부딪치게 된다. 물론 2개의 산업이 서로 포함관계에 있어 소프트웨어 산업 기반위에서 데이터베이스 산업이 개화하는 것이 가장 바람직한 것이지만 데이터베이스 산업은 예를 들어 “국가사회의 정보화” 등 그 범위가 한정된 특정 목적을 갖고 있기 때문에 목적과는 다른 결과를 빚을 수도 있는 것이다.

둘째로 데이터베이스의 생산구조에서 오는 문제점을 고찰해 보아야 한다. 생산구조란 상품의 특성과 함께 수요 시장의 구조로부터 판단해 볼 수 있다. 이 상품의 경우 생산측면에서 다품종 소량수요, 수요 측면에서 한정된 이용 계층으로 특징지어진다. 물론 국민 모두에게 필요한 보편적 수요를 갖고 있는 정보도 있지만 이러한 정보는 공기나 물처럼 이용자들이 간과하는 경향이 있다. 더욱이 이 경우에는 국민 전반의 문화나 의식구조가 크게 영향을 미친다.

따라서 대량수요가 발생하기는 어렵고 이것이

생산의 영세성으로 그대로 영향을 준다. 또 다른 한편으로는 대부분의 서비스 산업이 이용자와 서비스 제공자로 구성되는데 비하여 데이터베이스 사업은 원시자료 보유자, 자료 수집자, 자료의 가공을 통한 정보 생산자, 정보축적과 정보 유통자 등 낯설고 영성한 새로운 사업자의 등장으로 사업 구조가 훨씬 복잡하다. 생산의 영세성과 사업 구조의 복잡성은 데이터베이스 생산을 모험사업으로 만드는 주된 원인이 된다.

따라서 데이터베이스 생산은 주로 시장의 안정성과 사업의 복잡성을 피하는 방향의 in-house 데이터베이스, 예를 들어 한국 통신의 전화번호 데이터베이스, LG-EDS나 SDS 등 재벌 그룹내의 전산화 사업을 근간으로 시작되고 있다. '96 데이터베이스 백서 표 2-2-3, 2-2-5, 2-2-6을 참조하여 '95년말 현재 우리나라의 데이터베이스 생산 기관을 보면 이러한 사실이 극명히 드러난다.

- 총 438개 생산기관중 DB 전모제작기관은 74개업체 (16%)에 불과하다
- 자본금 규모를 보면 1억 미만이 89개 업체 (20.3%)인데 비하여 100억 이상이 64개 업체 (14.6%)에 이르고 있으며
- 종업원 규모로 보면 10명 미만이 82개 업체 (18.8%)인데 비하여 1,000명 이상이 64개 업체(14.6%)에 달하고 있어 데이터베이스 생산업체의 규모가 지극히 양극화하고 있다.

셋째로 데이터베이스 생산능력에서 오는 문제점을 살펴볼 수 없다. 여기에는 기술, 특히 변화가 빠른 기술 문제와 이 변화에 대응하는 표준 문제, 그리고 수공업적 소량생산인가? 자동화된 대량생산인가를 살펴야 하는 생산 양식이 있다. 데이터베이스란 하드웨어 산업, 소프트웨어 산업이 망라된 기술의 집약체이다. 그중에서도 운영시스템(OS 등) 등의 소프트웨어 산업과 데이터베이스 관리시스템(DBMS) 등의 미들웨어

(Middle ware) 산업을 기본으로 한다. 그러나 불행하게도 우리 나라에는 이들 기술이 특히 상업화된 기술이 거의 없다. 따라서 대부분의 데이터베이스 생산이 외제기술의 국내 적용이라는 차원에서 이루어지고 있어 국내에 알맞은 생산의 효율성이나 경제성을 확보하는데 어려움이 있다.

이런 환경에서 표준화란 기술의 표준화라기보다는 사업 모형의 개발 수준을 넘지 못하고 있으며 이것은 비디오텍스, 인터넷 멀티미디어로 하루가 다르게 변하고 있는 정보 상품을 체계적으로 발전시켜가는데 어려움을 주고 있다. 더욱이 정보시대의 도래가 정보의 생활화, 정보의 대량생산, 대량소비라는 패턴(pattern) 위에서 이루어진다고 볼 때 이런 정보의 개발을 저해하는 각종 문화적 요소 예를 들어 비공개성과 함께 기술, 표준 그리고 정보의 소비문화는 우리의 생산능력을 취약하게 하는 근본 문제가 되고 있다. 다행인 것은 우리의 문자인 한글, 우리의 언어인 한국어를 컴퓨터가 알아 듣고 처리할 수 있는 연구 즉, 우리글이나 우리말의 정보자원화에 대한 연구가 생산 현장이나 연구소, 학계 등에서 활발하게 진행되어 궁극적으로 21C에는 한민족도 정보시대의 중심적 기능을 수행할 준비를 갖추어 나가고 있다는 점이다.

마지막으로 데이터베이스 유통형식에서 생산에 미치는 영향을 살펴볼 필요가 있다. 여기서는 공중(public)을 대상으로 하느냐의 여부와 서비스 제공 방식을 ON-Line Real time으로 하느냐의 문제가 생산에 어떤 영향을 미치는가의 문제다. 데이터베이스는 국민 일반을 대상으로 하는 것(Open User Group)과 제한된 사람을 대상으로 하는 것(Closed User Group)으로 구분된다. 전자는 비교적 자료의 변경이 많지 않고 대량생산이 뒷받침될 수 있는 특징이 있는데 비하여, 후자는 자료의 변경이 많고 소량 생산의 고품인 경우가 많다. 후자의 경우 또 하나의 범주는

자가 사업 목적으로 개발한 데이터베이스로 자가라는 사업범위의 광, 협에 따라 사업의 안정성에 직접 영향을 준다.

예를 들어 LG-EDS가 LG-Group내의 전산화 사업을 추진하는 경우 안정된 생산기반을 갖출 수 있지만 중소기업에서 급여관리, 인사관리 프로그램을 사용하는 경우는 외주 등에 의해 영세, 불안정사업이 진행되는 것이다. 보급 형식의 또 다른 면에서 보면 필요한 정보를 CD-Rom 등의 형태로 수집 축적해 두고 필요할 때 간헐적으로 꺼내보는 오프라인형식의 데이터베이스와, 축적해두고 있는 데이터베이스를 네트워크(Network)를 통하여 항상 접속하여 인터랙티브하게 사용할 수 있는 온라인 리얼타임 형식의 데이터베이스가 있다. 미래 사회의 특징이 네트워크 사회, interactive한 사회, cyber사회라고 하는 것은 후자의 특성을 강조하고 있는 말이다. 이런 측면에서는 정보 사회의 실현은 통신 기술, 정보처리 기술, 정보생산 기술 등 첨단 기술의 총체적 구현으로 가능해지며 대규모 투자가 요청되는 것이다. 그러나 미래 사회를 성격지우는 가장 큰 요소로 작용하는 온라인 형식의 public용 데이터베이스의 생산은 그 중요성만큼 생산 Resource의 배분이 어렵다는 특징을 지니고 있다.

이상 데이터베이스 생산과 관련하여 산업 분류상의 위치, 생산구조, 생산능력, 유통형식 등을 개략적으로 검토하였다. 결국 정보는 기술집약, 지식집약의 특성을 지니고 있다. 그리고 그 특성상 통신사업, 신문, 방송 등의 언론사업, 생활 전반에 걸치는 부가가치 산업을 통합하는 기능이 있는 미래 지향적이고 아직은 리스크(Kisk)가 큰 산업분야이다. 그리고 시장에서 자생하는 사업이지만 정보화 시대의 조기 진입이라는 명제를 안고 있는 우리의 입장에서는 자생을 기다리기보다는 정책적 이니셔티브가 필요한 사업이다. 그렇

게 함으로써 이용자에게 편리한 생활을 제공할 수 있음은 물론 우리 나라의 산업구조 재편성이 이루어지고 우리도 21C의 중심 국가로 자리잡는다는 더 큰 목표를 실현할 수 있기 때문이다.

특히 초고속정보통신망이 완성되었을 때 일어날 수 있는 정보처리의 효율화, 대량 수요에 대응하는 Content 개발 및 생산도구 확보 등은 먼 훗날의 이야기라기보다는 오늘의 대응이 요구되는 부분이기도 하다. 실제 우리에게서 상용화된 데이터베이스 관리시스템(DBMS)도 없고 소프트웨어를 자동으로 생산하는 케이스 툴이나 다른 저작도구도 빈약하다. 더욱이 패키지 개발되어 있는 대형 프로그램은 많아도 이를 부품단위로 생산하여 활용하는 기술도 태부족이다. 그리고 현재 운용되고 있는 수작업으로 만들어 놓은 이 많은 프로그램이 유지보수를 요구할 때 대응하는 방법도 아직은 역부족이다. 산업계, 학계, 연구계는 중지를 모아야 할 것이고 정책부서에서는 목표를 향한 효율적인 정책개발이 이루어져야 할 것이다. DC