

데이터베이스 개발 비용산정에 관한 연구

신뢰성 있는 원가 산정과 기준 설정이 과제
...비용산정에 대한 모델 케이스 선정부터

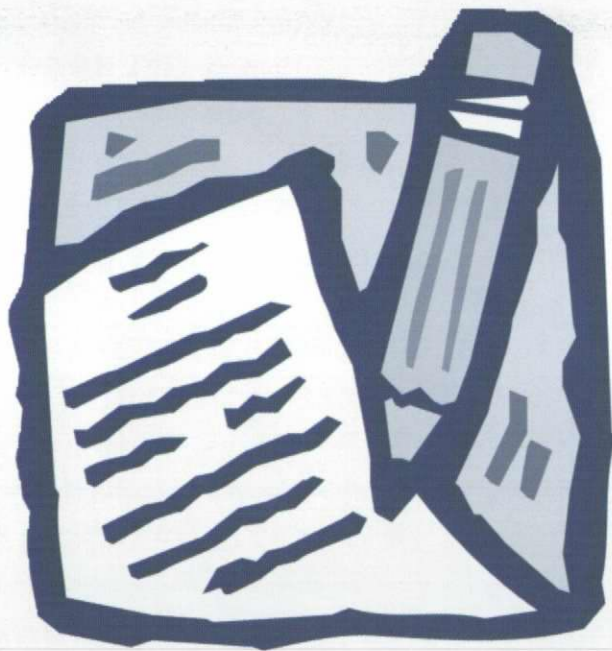
데이터베이스에 대한 논의 중 비용 산정에 대한 연구는 그리 큰 성과를 얻지 못했다. 그동안 논의는 많이 되었으나 실질적으로 데이터베이스 비용산정에 대해 집중적으로 연구한 것은 미비하다고 할 수 있다.

본고는 그러한 데이터베이스 비용산정에 대한 방향 설정과 이후 연구 과제에 대해 기술하고 있어서 향후 데이터베이스 비용산정에 기초 자료로 제공되기를 기대한다.

-편집자 주-

글 / 함정훈

한국데이터베이스진흥센터 연구원



1. 서론

급속히 늘어나는 정보 수요를 충족시키기 위해 데이터베이스 제공업자 및 정부는 최근 몇 년 동안 데이터베이스 산업육성에 힘을 기울였다. 그 일환으로 데이터베이스진흥정책이 수립되어 그 중요성 인식과 보급에 괄목할 만한 성장을 거두게 되었다.

그러나, 데이터베이스 산업의 급격한 성장에도 불구하고 데이터베이스에 대한 가치 인식 미비, 장기적인 안목에서의 과감하고 적극적인 기술 투자 미흡, 전문 기술인력의 만성적인 부족, 국내 시장 협소, 데이터베이스 개발업체의 영세성 등 질적인 수준은 데이터베이스 산업의 가시적인 성장에 미치지 못하고 있다.

특히, 데이터베이스 개발을 발주 또는 수주 할 경우 객관적이고 합리적인 데이터베이스 개발비용 산정 기준과 연구 결과가 부족하여 데이터베이스 개발자들이 노력한 만큼의 정당한 가치를 인정받지 못하게 되는 어려움을 겪고 있다. 이러한 개발자의 어려움은 개발업체의 채산성 및 재무구조를 악화시키고, 결국 데이터베이스 품질과 연결되어 악순환을 거듭하게 되는 것이다.

이에 정부는 1989년 4월 과학기술처 고시로 『소프트웨어 개

발비용 산정기준』을 발표한 바 있다. 그러나, 데이터베이스 개발과 소프트웨어 개발은 여러 측면에서 많은 차이를 가지고 있기 때문에 데이터베이스 개발의 경우 『소프트웨어 개발비용 산정기준』을 채택하지 못하는 것은 자명하다.

데이터베이스와 일반 소프트웨어와의 개발 상의 차이점을 비교해 보면 <표1-1>과 같다.

<표1-1>에서 보는 바와 같이 데이터베이스 개발은 일반 소프트웨어 개발과는 그 개발 과정 및 개발 요소 등에서 큰 차이가 있다.

이러한 차이에도 불구하고 소프트웨어 개발비용 산정 기준을 데이터베이스 개발에 적용되어 왔으며, 그 결과 타당성이 인정될 수 없는 개발비용이 산정되어 왔다.

이에 1994년 한국데이터베이스진흥센터에서 『데이터베이스 구축비용 산정기준 및 서비스 이용 요금 체계 설정을 위한 기초조사』 연구를 실시하였으며, 정보 산업의 네트워크화, 오픈시스템화, 멀티미디어화, 다운사이징화 등의 추세에 따라 SI체제로의 정보시스템의 구축 프로젝트가 일반화되면서, 97년 정보통신부 고시 제1997-57호로 『소프트웨어 사업 대가 기준』에 데이터베이스 개발비용 산정기준이 포함되어 고시되었다.

그러나 정보기술의 급속한 발전, 사회 각 부문의 폭발적인 정보화 요구의 증가로 데이터베이스 개발 사업의 복잡화, 대규모화가 이루어지고 있는 환경에서 기존 『소프트웨어 사업 대가 기준』으로는 성공적인 데이터베이스 개발, 운영을 실행하기 어려운 상황이다.

본 연구에서는 이러한 국내의 환경을 반영하는 효과적인 데이터베이스 개발비용 산정 모델을 만들기 위해 COCOMO II, 수학적 모델을 참고하였으며 기본적인 공수 및 비용산정은 소프트웨어 사업 대가 기준을 참고하였다.

본 연구의 구성은 제 2장에서 모델개발을 위한 다양한 국내의 모델을 조사해 보고, 제 3장에서는 데이터베이스 개발비용 산정 모델을 제시한다.

그리고 마지막으로 제 4장에서는 향후 연구방향을 제시한다.

2. 비용산정 기준의 이론적 고찰

2.1 알고리즘 모델(Algorithm Model)

특정 데이터베이스 척도와 프로젝트 비용 정보를 바탕으로 개발된 모델로 데이터베이스 개발비용에 영향을 미치는 중요한 변수 한 개 혹은 여러 개의 수식에 대입하여 비용을 추정하는 방법이다.

2.2 전문가의 판단(Expert Judgement)

데이터베이스 개발에 사용하는 기술에 대한 여러 전문가와 데이터베이스를 이용할 분야의 전문가 판단에 의한다. 이들 각각은 나름대로 비용을 산정하고 최종적인 비용산정은 전문가 합의에 의해 결정한다. 개발 초기 단계에서 과거 유사 프로젝트의 규모 관련 자료가 거의 없거나 불충분 할 경우, 규모산정에 경험 있는 전문가 다수가 참여하여 거시적 차원에서 규모를 산정하는 기법이다. 이 기법은 델파이 기법의 일종으로 이론적인 배경은 단일 수준의 전체 규모산정 보다 다수 예측의 합이 더 정확하다는 것을 전제로 한다. 이 기법의 적용 절차는 먼저 델파이 그룹을 구성하고 네 차례에 걸친 응답결과로부터 중간 값 또는 제2사 분위수를 최종 산정값으로 선택하게 된다. 델파이 그룹에 속한 6-8명을 그 분야의 전문가로 지정하고 문제를 명확히 정의한 다음 이전에 유사한 산정경험을 고려하여 예측 값을 제시하도록 한다. 세부적인 절차는 4라운드로 구성되는데 산정자는 매 라운드마다 최소값, 최대값, 제1, 2, 3 사분위수, 중간값, 평균값에 대한 근거를 제시하여야 한다.

2.3 하향식 산정(Top-down Estimation) 및 상향식 산정(Bottom-up Estimation)

하향식 비용산정은 제품의 전체적인 특성을 고려하여 결정되며, 이 비용은 각 구성 부분에 할당된다. 상향식 비용산정은 각 구성 부분의 비용이 산정되고 이 비용이 전체적으로 더해져서 최종적인 비용을 산정하는 방법이다.

2.4 점수 모델

점수에 의해 규모를 산정하는 모델은 소프트웨

<표1-1> 데이터베이스 개발과 소프트웨어 개발의 비교

항목	구분	데이터베이스	소프트웨어
핵심요소		데이터	프로그램
프로그램 크기		(DBMS) + 소규모	개발 업무의 규모와 비례
사용언어		SQL, 4GL, C	4GL, C, COBOL 등
개발인력의 주된 소요점		데이터 처리	프로그램 개발

주) 데이터베이스진흥센터, 『데이터베이스 구축비용 산정기준 및 서비스이용 요금체계 설정을 위한 기초조사』, 1994

어에서 기존의 프로그램 스텝수와는 전혀 다른 척도인 소프트웨어 기능량(Function Point)을 정량화하여, 소프트웨어규모를 나타내는 것이다.

기능점수(Function Point) 모델이란 생산성 측정기법으로 사용언어, 작성자의 기술, 개발환경, 4GL 등과 같은 시스템 작성 방법과는 직접 관계가 없고, 일관된 척도로 시스템의 가치를 평가하는 것이다. 이렇게 함으로써 종래의 프로그램 수나 소스코드(스텝수)에 의한 접근의 한계를 극복할 수 있다. 완료된 프로젝트나 가동중 시스템의 기능점수 실적치를 축적하여 개발공수와 관계를 분석하며, 견적에도 응용 가능하게 된다. 기능점수법은 미국 IFPUG(International Function Point User Group)을 중심으로 개발이 계속되고 있으며, ISO의 새로운 검토항목을 인정되었다.

2.5 기타

유추에 의한 산정(Estimation by Analogy)방법은 동일한 분야에서 상이한 프로젝트의 개발이 끝났을 때 적용할 수 있다. 새로운 프로젝트 비용은 이미 완성된 프로젝트로부터 유추하여 산정한다.

Parkinson's Law(파킨스의 법칙)이란 작업은 가용한 시간까지 계획된다는 의미이다. 이것은 소프트웨어 비용 측면에 적용한다면 프로젝트에 쓸 돈이 있으면 다 쓰게 된다는 것이다.

능력별 지불(Pricing to Win)이란 데이터베이스 비용은 고객이 프로젝트에 대해 지불할 수 있는 비용 그 자체로 산정된다는 것이다. 합리적인 가격 산출 기준이나 가격 예측의 기법을 잘 알지 못하는 상태에서 개발비용을 산정하는 방법이다.

위에서 제시한 각각의 비용산정 기법들은 장단점을 가지고 있으며, 상호 배타적이기보다는 상호 보완적으로 적용된다.

3. 데이터베이스 개발비용 모델 제안

3.1 데이터베이스 정의

데이터베이스 개발비용 산정 기준의 제안을 위하여 사용되는 용어는 다음과 같이 정의하며, 표현되는 단어는 일반적으로 사용되는 의미와 상이(相異)할 수 있다.

$$Y = \sum_{i=0}^n S_i + \sum_{j=1}^m D_j + \sum_{k=0}^l DE_k + d$$

Y = 데이터베이스개발 비용
 S_i = 소프트웨어개발 비용
 DE_k = 원시자료, DBMS 등 각종 경비
 D_j = 데이터제작 비용
 d = 제경비

가. 데이터베이스 개발이란 데이터베이스 구축작업과 이에 필요한 소프트웨어 개발작업을 통하여 하나의 유용한 시스템을 개발하는 작업과정을 말한다.

나. 데이터베이스 구축이란 사회적으로 산재하고 있는 방대한 원시자료를 수집·분석·분류하여 이용자에게 유용하고 의미 있는 정보로 설계하여 수집한 자료를 가공·제작·제공하는 일련의 작업과정을 말한다.

다. 원시자료 수집이라 다양한 형태, 방대한 양으로 존재하는 정보원천 중 이용자에게 유용한 정보로 전환될 수 있는 재료를 선별하여 수집하는 작업과정이다.

$$\text{데이터 제작비} (\sum_{j=1}^m D_j) = D_1(MmC \times q) + D_2(MmC \times q) + D_3 \dots$$

D_j = 데이터제작 비용
 MmC = 미디어별 데이터제작 단가
 q = 수량
 j = 데이터제작 미디어 종류

라. 데이터 제작이란 수집된 원시자료를 일련의 체계화된 작업절차에 따라 이용자에게 유용한 의미 있는 정보형태로 전환하여 표현하는 작업과정을 말한다.

마. "스트로크"라 함은 자료입력원이 자료입력 목적수행을 위하여 전자계산기 자판의 각종 키들 중 어느 하나를 한 번 누르는 것을 말한다.

3.2 데이터베이스 개발 모델 제안

〈표2-1〉 데이터제작 종류 및 기준

구분	종류	기준		
		단위	기준 양	
데이터제작	문자	숫자	스트로크	1,000,000 (스트로크)
		영어		
		한글		
		한자		
		HTML		
		XML		
	기타(SGML...)			
	이미지	흑백	장	100장
		컬러		
	그래픽	2D	장	100장
3D				
애니메이션		초	10초	
동화상		초	10초	
음성·음향·음악		초	1,000초	

제작에 투입되는 공수는 제작하는 데이터 난이도에 따라 구분되며 <표2-2>와 같다. 이 기준은 현행 소프트웨어 대가기준의 데이터제작비 기준을 재구성한 것으로 공정 및 그에 따른 투입공수가 현황 및 상황에 따라 변경될 수 있다.

데이터의 난이도를 평가하기 위해 평가를 위한 객관적인 평가 척도 및 평가기준, 절차 등에 대한 연구가 이루어져야 한다. 예컨대 이러한 평가는 결국 데이터베이스 품질과 밀접한 관계가 있어, 평가에 기초자료로 활용될 수 있다.

<표2-2>는 데이터 미디어별 공정 및 소요공수를 현행 『소프트웨어 사업대가 기준』을 참고하여 재구성하였다. 이러한 공정별 공수표는 사업을 발주하는 업체가 그간의 노하우(Knowhow)를 바탕으로 시장현황을 고려하여 수주자 및 평가자로부터 타당성을 검증 받아야 한다.

또한 위에서 보면 난이도를 단순, 보통, 복잡으로 구분하여 소요공수를 산정한 것을 볼 수 있다.

이러한 구분을 위해서는 평가척도, 기준 및 평가방법이 필요하다. 평가척도는 일반적으로 5점 척도를 사용한다. 평가기준은

$$\sum_{k=0}^l DE_k$$

DE ₀ : 원시자료구입비	DE ₁ : DBMS구입비
DE ₂ : 메뉴화면 및 웹사이트 제작비	DE : 기타 개발비

공정별 요구난이도, 자료가공도, 공정복잡도, 숙련요구도, 요구구현도, 산출물완성도 등으로 구분하여 각각을 평가할 수 있다. 하지만 평가기준이 평가주체, 평가방법에 따라 차이가 있을 수 있다는 문제가 있다.

뿐만 아니라 데이터 미디어별로 평가하는 평가 항목들이 차이가 있으므로 미디어별로 평가항목을 개발에 대한 많은 연구가 필요하다.

데이터제작에서 자료입력원들은 시간이 지남에 따라 교육과 반복 효과로 인하여 작업은 더욱 능률적이고 생산성 또한 높아진다. 따라서 학습효과에 의해 제작되는 데이터 규모가 클수록 제작 단위당 투입공수를 줄일 수 있다.

하지만 규모별 보정계수에 사용할 학습률 데이터가 충분하지 않으므로 모델에서는 학습효과를 고려하지 않기로 한다. 학습효과에 의해 줄어드는 소요공수는 일정기간 평균 소요공수를 적용함으로써 어느 정도 근접할 수 있다.

원시자료구입비, DBMS 구입비 등의 각종경비(DE_k)은 데이터베이스 개발을 위한 소프트웨어개발 및 데이터제작 비용을 제

외한 모든 개발비 및 구입비를 포함한다.

원시자료 구입비는 사회적으로 그 수집경로, 수집방법, 수집형태, 수집대가 등이 매우 다양하고 수집된 원시자료가 실제 제 공정보로 제작되어 서비스되는 비율도 서로 상이하다. 또한 수집자료에 대한 저작권, 품질 등에 대한 기준이 모호하여 일률적인 단가를 정하여 대가를 산정하는 것이 사실상 불가능하다. 따라서, 각 사례별로 현재 데이터 수집의 시장현황 및 경험에 따라 총량을 기준으로 비용을 산정한다.

DBMS를 직접 개발하여 사용하는 경우도 있지만 일반적으로 시장에서 DBMS를 구입하여 사용한다. 따라서 DBMS는 시장 가격에 따라 구입한 비용을 산정한다.

메뉴화면 제작비는 메뉴화면의 바탕화면이나 안내정보 제작으로 데이터제작비에 의해 산정한다. 또한 서비스를 위한 웹사이트 제작이 필요한 경우 일괄계약이나 소프트웨어, 데이터제작비 기준에 의해 산정할 수 있다.

제경비는 출장비를 포함한 장소임대비, 장비구입비, 인쇄비, 자문비 등으로 일반적으로 개발비에 110%을 일률적용하고 있다.

이렇게 적용하는 것에 대해 일부 개선에 대한 논의가 있지만, 시장이 이런 비용을 인정하고 있어 110%를 적용하기로 한다. 하지만 제경비는 시장현황에 따라 변동될 수 있다.

4. 결론 및 향후 연구방향

정보산업이 국가의 핵심산업으로 성장해 감에 따라 정보산업의 기반을 이루는 데이터베이스 산업의 중요성이 증대되고 있다. 그러나 정당한 데이터베이스 구축 사업비 책정 및 대가의 지불이 미흡하여 데이터베이스 개발업체의 체산성 및 재무구조는 악화되고 이로 인해 데이터베이스의 품질은 저하되어 발주자의 신뢰성이 떨어지고 있다.

이러한 문제는 신뢰할 만한 데이터베이스 구축비용 산정기준이 부족하기 때문이다. 따라서 신뢰할 만한 기준 및 원가산정 등의 연구가 지속적으로 필요하다. 더욱이 현재 정확한 비용산정 기준을 위하여 반드시 과거 개발된 데이터베이스 비용에 대한 자료를 DB화하여 비용산정에 사용할 수 있도록 하는 작업이 선결 과제이다. ☞