

데이터베이스 품질 평가를 위한 모형 개발*

-텍스트 데이터베이스 내용을 중심으로-

Developing a Model for Quality Evaluation of Text Database Contents

장혜란(Hye-Rhan Chang)**

목 차

1 서 론	4.1 지표별 평가요소
2 데이터베이스 평가의 발전	4.2 측정방법
3 내용 평가지표의 설정	4.3 평가척도의 개발
4 평가요소와 측정방법	5 결론 및 제언

초 록

데이터베이스의 품질 평가에 관한 선행연구 결과를 종합적으로 검토하고 텍스트데이터베이스를 중심으로 내용평가를 위한 모형을 개발하였다. 구체적으로 5개 지표의 16개 평가요소를 설정하고, 각 요소별 평가 데이터의 수집과 측정 방법, 그리고 평가척도를 제시하였다. 평가의 실행과 추진방안을 제언하였다.

ABSTRACT

Based on the results of previous evaluation efforts, a database quality evaluation model, applicable to text databases, is developed. Focusing on database contents, 5 evaluation criteria consisting of 16 elements are defined. For each element, data collection method along with measuring process is established. And an evaluation scales are also provided. The concluding section suggests several areas for implementation and future development.

키워드 : 데이터베이스, 품질평가, 평가모형, 평가지표, 평가척도

* 이 논문은 2000학년도 상명대학교 사회과학연구소 연구비의 지원을 받아 수행되었음.

** 상명대학교 문헌정보학과 부교수

■ 논문 접수일 : 2000년 11월 15일

1 서 론

컴퓨터와 통신 기술의 발전으로 데이터베이스 생산과 유통이 급격하게 증가하였으며, 데이터베이스 탐색은 도서관 및 정보서비스기관은 물론 일반 대중에게도 확산되었다. 데이터베이스에 대한 사회전반의 요구는 양적인 발전을 벗어나 질적인 발전을 추구하고 있으며, 여기서 우리 분야가 당면하게 된 문제는 데이터베이스 품질평가에 관한 것이다.

특히 우리나라의 경우에는 정부의 정보화정책에 힘입어 데이터베이스가 양산되었는데, 1992년 이후 연평균 42.9%의 비율로 증가하여 1999년 12월 현재 상업적 혹은 공공적으로 이용가능한 데이터베이스는 3,290종에 이른다(한국데이터베이스진흥센터 1999). 데이터베이스의 급성장과 더불어 평가의 필요성이 끊임없이 제기되고 있다. 컴퓨터와 네트워크의 기반은 조성되어 있으나 실제로 유통될 정보의 내용이 부실하다는 것이다.

우리나라 데이터베이스의 가장 큰 문제는 그 내용에 있다고 할 수 있다. 그 이유는 본질적으로는 데이터베이스가 상품으로 개발되어 자체적 품질관리가 이루어졌던 선진국들과 달리, 서지와 참고자료가 발달되지 못한 상황에서 짧은 기간 동안 급속히 제작 유통됨으로써 일정수준의 품질에 도달할 수 없었기 때문일 것이다.

상품과 서비스에 대한 품질평가는 그 존재이유가 되는 목적과 목표를 어느 정도 달성하고 있는가를 나타내 주는 중요한 정보를 제공함으로써 이용자들의 선택을 결정하고 해당 산업의 발전에 영향을 미친다.

이와 같은 품질 평가의 중요성에도 불구하고 데이터베이스 부문의 평가가 수행되지 못하고

있는 이유는 일차적으로 객관적 평가를 가능케 하는 정립된 모형이 없기 때문이다. 결과적으로 질적으로 낮은 데이터베이스가 유통되고 있으며, 모든 탐색자들은 개별적으로 품질을 판단해야 하는 부담을 지니고 있다.

본 연구의 목적은 데이터베이스 품질 평가에 관한 선행연구를 종합적으로 조사하고 검토하여 적용가능한 평가모형을 개발하는 데 있다. 데이터베이스 산업이 발전됨에 따라 다양한 유형의 데이터베이스가 생산되고 있으나, 본 연구는 상용 온라인 데이터베이스의 가장 많은 부분을 차지하고 있는 텍스트 데이터베이스로 범위를 한정한다. 데이터베이스의 품질은 컴퓨터와 통신기술 그리고 관련서비스와 밀접하게 연결되어 있지만, 데이터베이스 상품의 핵심이 되는 내용 평가를 주요대상으로 한다.

데이터베이스의 품질을 평가하기 위한 모형은 단지 기준의 명세서와 같은 것은 아니다. 품질 평가는 수준을 결정하는 과정이며, 일정한 규칙에 따라 데이터를 수집하여 수치를 부여하는 측정이 필수적이다. 따라서 본 연구에서는 데이터베이스의 내용을 평가할 지표를 설정하고, 각 지표별 측정요소와 측정방법을 제시하고 종합점수를 산정할 수 있는 평가모형을 개발하고자 한다.

이와 같은 평가 모형의 개발은 데이터베이스의 평가가 요구되고 있는 현황에 비추어 지극히 필요한 것이며, 데이터베이스 제작자에게는 품질진단 및 개선을 가능케 하는 지침을 제공할 수 있고, 탐색전문가와 최종이용자에게 데이터베이스 선택에 필요한 정보를 제공할 수 있는 토대를 마련함으로써, 품질개선과 정보이용 활성화에 기여할 것이다.

2 데이터베이스 평가의 발전

데이터베이스 품질에 관한 문제는 여러 관련 분야의 학자들에 의하여 논의되어 왔다. 그 중 텍스트 데이터베이스의 평가는 전통적으로 문헌정보학 분야에서 이루어졌다. 초기의 객관적인 평가는 인쇄본 이차자료 서비스를 대상으로 수행되기 시작하였으며, 기계가독형 데이터베이스가 출현하고 온라인화 되면서 그 편리성으로 인해 평가연구가 더욱 활발해 졌다. CD-ROM 데이터베이스는 반복탐색과 데이터베이스 사용자 문제를 해결함으로써 평가연구의 새로운 전환점이 되었으며, 1990년대에 들어와 인터넷을 활용한 탐색이 증가되면서 여러 가지의 평가 모형 개발이 시도되었다.

이차자료의 품질에 대한 실증적인 연구는 1960년대부터 시작되었다. 마틴과 슬레이터는 포괄성에 대한 고전적인 연구를 수행하였으며 (Martyn and Slater 1964), 이후 이차자료의 포괄성, 중복성, 유일성에 관한 다수의 연구가 수행되었다. 이차자료의 성능 평가에는 레코드 분석 이외에 다양한 방법들이 사용되었는데, 그 중에는 이용자들의 견해를 조사하기 위한 면접 및 설문조사도 적용되었다 (Cluley 1968, Drage 1969, Hall 1972). 한편 적합성, 일반성, 사용용이성, 최신성 등의 요소가 색인의 질을 평가하기 위한 기준으로 언급되기도 하였다 (Azgaldov 1969).

1970년대에 들어와서는 유사한 분야, 주제, 혹은 자료유형에 따라 다수 데이터베이스를 상호 비교하는 연구들이 수행되었는데, 그 중 베어먼과 쿤버거의 연구는 대규모의 것으로 데이터베이스 14종의 포괄성을 평가하였다 (Bearman and Kunberger 1977).

한편 데이터베이스 품질에 관한 관심은 정확성에 대한 측정과 연구에도 집중되었다. 초기에는 철자에러에 대한 검출과 수정이 중심이었으나, 점차 특정 유형의 에러를 탐지하거나(예, 중복레코드, 불완전한 레코드) 전거파일에 의한 정정으로 발전하였으며, 데이터베이스 생산자들이나 서지유틸리티서비스의 노력 그리고 관련소프트웨어의 개발에 힘입어 상당 수준의 품질을 성취할 수 있었다(O'Neill and Viziner-Goetz 1988).

1980년대 후반에 기존의 데이터베이스 중 이용도가 높은 것이 CD-ROM 버전으로 출시되기 시작하였는데, CD-ROM의 이용은 평가에 대한 하나의 전환점이 되었다.

Herther(1988)는 CD-ROM 데이터베이스 평가를 위한 8가지 기준을 제시하였는데, 그 내용은 정확성과 다큐멘테이션, 적합성, 이용편의성, 권위와 신뢰성, 내용, 유사데이터베이스와의 비교, 갱신, 가치 등이다. Foulds와 Foulds(1990)는 이것을 보다 상세한 체크리스트로 발전시켰다. CD-ROM 데이터베이스에 대한 평가는 기존의 평가요소와 더불어 사용자인터페이스, 도서관장서로서의 가치 등을 포함하게 되었다. 사용자인터페이스는 최종이용자 탐색이 확산되면서 그 중요성이 부각되기 시작하였으며, 인터페이스를 평가하기 위한 다수의 기준이 제시되었다(Harris 1990, Large 1991, Zink 1991).

정보상품으로서의 데이터베이스의 품질과 가치에 대한 관심의 확산은 1990년대 초에 가시화되기 시작하였으며, 이러한 경향은 90년대를 통하여 지속되었다.

1989년 Finnish Society for Information Service는 데이터베이스 품질에 대한 개념을 정

의하고 평가기준을 만들기 위한 연구에 착수하였다. 이 연구는 핀란드 교육부의 지원을 받아 진행되었는데, 연구집단을 결성하여 품질기준을 마련하는 데 1년 이상이 걸렸고, 기준의 적용 가능성을 테스트하는 데 7개월이 소요되었다. 이 기준은 시스템과 통신연결, 탐색언어와 기술, 탐색소프트웨어의 효과, 데이터베이스 내용, 검색보조, 비용 등 총 6개 부문에 걸친 109개의 문항으로 구성되어 있다(Juntunen et al. 1991).

거의 같은 시기에 미국에서는 SCOUG (Southern California Online User Group)가 데이터베이스의 품질을 평가하기 위한 기준을 개발하였다. 그 내용은 일관성, 범위와 포괄성, 최신성, 정확성, 접근가능성, 완전성, 출력형태, 다큐멘테이션, 이용자교육과 지원, 비용대가치 등 10개 부문으로 구성되어 있으며, 각 부문은 또한 다수의 문항 리스트를 포함하고 있다 (Basch 1992).

SCOUG와 Finnish Society의 평가기준은 모두 온라인 데이터베이스의 내용과 서비스의 다양한 양상을 광범위하게 포함하고 있으나, 핀란드 기준은 서지데이터베이스에 대한 정보전문가의 판단에 초점을 둔 반면에 SCOUG의 기준은 서지, 본문, 디렉토리 등 텍스트 데이터베이스 전반을 대상으로 최종이용자의 판단에 초점을 두고 있다.

위의 두 가지 평가기준은 데이터베이스 평가에 대한 관심을 불러일으키고 차후의 평가연구에 지대한 영향을 미쳤다. 그러나 부문별 세부 사항을 나열식으로 제시하고 있으며, 구체적인 평가방법이나 표준이 나타나 있지 않고, 데이터베이스와 탐색서비스 그리고 이용자 만족에 관한 사항이 명확하게 구별되어 있지 못한 제한점

을 가지고 있다.

SCOUG에서 개발한 기준은 영국의 CIQM(The Center for Information Quality Management)에서 데이터베이스를 등록하는데 적용되었다. 1993년 영국의 LA(The Library Association)와 UKOLUG(United Kingdom OnLine User Group)은 데이터베이스 품질에 대한 일종의 클리어링하우스 활동을 위하여 CIQM을 창설하였다. CIQM은 데이터베이스 제작자로부터 데이터베이스 현황에 대한 명세서를 받아 공표함으로써, 공표하지 않은 데이터베이스와 차별을 시도하였다. 명세서의 내용은 SCOUG기준과 유사한 9개 부문으로 구성되어 있고, 그외에 품질관리에 대한 진술, 통계치, 레코드의 필드구성등을 포함한다. CIQM은 등록된 데이터베이스에 대한 디렉토리를 매년 2회 생산하며, 1996년부터는 비정기 web-newsletter인 dbQual 을 간행하고 있다.

CD-ROM 데이터베이스 평가기준에 대한 Harry와 Oppenheim(1993)의 두 가지 연구는 데이터베이스의 품질을 점수화하고, 4개의 데이터베이스를 평가하여 등급화를 예시하였다는 점에서 종래의 연구와 큰 차이가 있다. 이들은 상품명세, 기술명세, 다큐멘테이션/지원, 데이터베이스, 이용자인터페이스, 탐색, 출력, 신뢰성 등 8가지 부문에 대하여, 부문별 세목을 정하고 세목별 점수와 부문별 평가자 부가점수를 합하여 데이터베이스 점수를 산정하였다.

90년대 후반에는 인터넷 데이터베이스, 특히 WWW데이터베이스의 평가에 대하여 관심이 집중되어 있다. 다수의 연구가 수행되었으며, Web 사이트 평가에 대한 서지도 나오고(Nicole 1998), 한편으로는 인터넷정보원에 대한 학문별, 혹은 주제별 평가사이트도 등장하였다.

WWW데이터베이스 평가의 기준으로 제시된 사항을 고찰해보면, 종래에 제시되었던 데이터베이스 평가기준에 추가하여 저자의 권위, 정보의 안전성, 객관성, 링크되는 자원의 품질 등을 기술하고 있으며 멀티미디어의 경우에는 디자인, 음향 등에 관한 사항이 추가되기도 한다. 제시된 평가기준들은 대부분 명확한 지침이 아닌 진술에 불과하며, 평가사이트는 유용한 정보를 제공하지 못하고 있다.

보다 최근에 미국의 대학도서관 콘서시엄의 하나인 HELEN에서는 참고사서가 중심이 되어 데이터베이스 평가를 위한 가이드라인을 마련하였는데, 그 내용은 색인된 정보원, 색인된 학술지, 포괄성, 색인, 탐색기능, 탐색지원, 스크린 설계, 본문/그래픽 내용, 출력, 신뢰성, 접근성, 가격 등 12개 부문의 54개 질문으로 구성되어 있다(Brennan et al. 1999).

우리나라에서도 데이터베이스 품질 평가기준에 대한 연구와 저술이 이루어졌다. 그 중에서 가장 포괄적으로 접근한 이국희(1998)의 연구에서는 데이터베이스의 품질을 내용과 서비스로 나누어 내용은 정확성, 완전성, 현행성, 일관성의 4부문으로 서비스는 검색성, 사용용이성, 사용자지원성의 3부문으로 구성하였다. 기준을 적용하기 위한 측정도구는 설문이며 대체로 사용자 만족도/인식에 기초를 두고 있다.

이상과 같이 데이터베이스를 평가하기 위한 연구와 노력이 지속되어 왔으나, 품질측정과 평가에 대한 많은 부분이 아직도 연구대상으로 남아 있다. 평가기준은 평가부문과 부문별 세부사항을 제시하는 데 불과하거나, 너무 많은 부분을 이용자인족에 의존하고 있으며, 측정결과에 대한 준거를 제시하지 못하고 있다. 데이터베이스제작자와 이용자들은 실제적인 평가 정보를

필요로 하고 있으나, 명확하고 구체적인 유용한 정보가 부족하고, 간혹 모호한 기준으로 등급을 매기고 있다(예, Cyberhound's). 신뢰성 있는 통보적인 결과를 제시할 더 나은 모형을 필요로 한다.

3 내용 평가지표의 설정

정보검색시스템은 축적된 정보, 하드웨어, 소프트웨어, 그리고 적합정보의 탐색과 디스플레이를 용이케하는 인터페이스로 구성되어 있다. 이 모든 요소들, 그리고 서비스 기관의 이용자 지원과 통신기술은 모두 탐색될 데이터베이스의 품질에 영향을 미친다. 이 요소들은 모두 필요한 것이며 이 중에 무엇이 중요한가의 견해는 연구자에 따라 상당한 논란을 초래할 것이다.

그러나 구성요소의 일부는(예, 통신, 하드웨어) 우리분야 이외에서 개발되고 있으며, 텍스트기반 탐색기능은 어느 정도 표준화되어 있고, 인터페이스의 개선은 하드웨어나(고속 프로세서, 마우스와 같은 입력장치) 소프트웨어의(윈도우기반OS, 그래픽) 개선과 밀접한 관계가 있다. 또한 데이터베이스 형태에 따라서 상호작용이 제한되어 있고(예, 온라인, CD-ROM vs. 인터넷) 이용자지원에 관한 사항도 명확한 책임의 구분을 가르기 어렵다. 정보시스템을 크게 들로 나누면, 축적된 정보와 정보에 대한 접근을 제공하기 위한 액세스메카니즘이 될 것이다. 따라서 데이터베이스 품질을 결정하는 요인 중에 일차적으로 고찰되어야 할 사항은 정보시스템의 기초가 되는 그 내용이다.

전술한 바와 같이 1990년 이래로, 데이터베이스 품질일반 및 내용평가에 대한 논의가 지속

〈표 1〉 선행연구에서 제시된 내용평가 지표

지 표	Finnish Society	SCOUG /CIQM	Harry & Oppenheim	HELEN	이국희
포괄성	○	○	○	○	
최신성	○	○	○		○
완전성	○		○	○	○
정확성	○	○			○
일관성	○	○			○
색 인	○			○	

되고 있다. 데이터베이스 내용을 평가하기 위한 여러 가지 지표가 제시되었는데, 각각의 기준은 제시한 지표와 구성요소에 대하여 다양한 견해를 나타내고 있지만, 내용의 품질을 평가할 지표가 무엇인가는 어느 정도 식별될 수 있다. 대표적인 연구를 중심으로 내용기준에 포함된 사항을 정리해 보면 〈표 1〉과 같다.

〈표 1〉에 나와 있는 바와 같이 평가기준에 제시된 지표는 연구마다 차이가 있으나, 데이터베이스의 질적인 평가지표가 될 공통적인 사항은 포괄성, 최신성, 완전성, 정확성, 일관성으로 드러났다.

포괄성은 목적하는 해당 정보를 데이터베이스가 실제로 포함하고 있는 정도를 나타내며, 데이터베이스 제작자의 정책에 크게 의존한다. 최신성은 데이터베이스 생산의 신속성을 나타내며, 신규레코드가 데이터베이스에 추가되는 시점이 시간적으로 적절인가 즉, 적시성과 관계가 있다. 데이터베이스 내용에 있어서의 완전성은 레코드를 구성하는 내용, 그리고 구성 필드의 값이 모두 기록되어 있는가를 나타내며, 이것은 탐색점 제공과 관계가 있다. 일관성은 레코드에 기입된 데이터표현이(예, 저자명, 약어, 출판일자) 동일한 정도를 말하며, 정확성은 데

이터 자체에 있는 에러정도를 나타내는데, freetext retrieval system 에서 색인은 레코드로부터 직접 생산되고, 검색은 전형적으로 완전 매칭에 의하므로 일관성과 정확성은 탐색에 영향을 미친다. 데이터베이스 내용의 품질을 논할 때 빈번하게 언급되는 것은 색인의 질이며, 정교한 색인의 통제에 중점을 두고 있다.

본 연구에서는 포괄성, 최신성, 완전성, 정확성, 일관성을 내용 평가의 지표로 설정하고 색인의 질은 레코드 구성 내용의 하나로 보아 완전성 지표에 포함시키코자 한다.

4 평가요소와 측정방법

4.1 지표별 평가요소

평가모형을 구축하기 위하여 설정된 다섯 가지 평가지표에 대한 범위와 내용을 결정하였다. 지표별 평가요소의 선택은 선행연구를 기초로 하였으며, 필수적인 요소에 한정시키코자 노력하였다.

포괄성은 기본적으로 포함되어야 할 정보가 데이터베이스에 어느 정도 포함되어 있는가를

나타내는 것으로, 레코드를 단위로 분석하여 %로 표시된다. 그밖에 간주되어야할 요소로는 수록 레코드 수의 적절성, 제시된 정보원(예, source list)의 적절성, 수록 정보자료 유형(예, 단행본, 정기간행물 등등)의 망라성, 소급정도의 적절성 등을 선택하였다.

최신성은 정보가 신속하게 데이터베이스에 적재되는 정도를 나타내며, 기본적으로 원정보의 생산일자와 해당 레코드가 입력된 일자와의 시차를 평가요소로 결정하였다. 데이터베이스 유형에 따라 갱신빈도의 적절성을 함께 간주하기로 정하였다.

완전성은 레코드 내용의 완전성을 나타내는 것으로, 레코드를 구성하는 필드의 적절성, 탐색가능 필드 지정의 적절성, 수록레코드 전체 중에 필드값이 모두 기록된 레코드의 비율, 그리고 색인과 관련된 어휘통제의 유무와 유지보완 정도를 평가요소로 선택하였다.

모든 데이터베이스는 에러를 포함하며, 이것은 원정보에서 비롯될 수도 있고, 수작업 입력 혹은 optical scanning입력에 기인하기도 한다. 에러는 업무의 효율성과 관계되며 컴퓨터 축적비용, 처리시간, 탐색효과에 치명적이다. 지난 20여년 동안 에러를 탐지하기 위한 기법이 개발 적용됨으로써 이 문제는 어느 정도 성공적으로 해결되고 있다. 정확성의 평가요소는 전체 레코드 중에 러가 없는 레코드의 비율로 결정하였다.

일관성은 레코드의 각 필드에 있는 내용이 일정한 형식을 갖는 정도를 나타낸다. 일관성 또한 탐색에 영향을 미치며, 정보표현에 있어서의 규칙의 적용과 관계가 있다. 구체적으로 기술상의 문제가 제기되고 있는 출판사항의 기술, 외국인 저자명 기술, 약어의 기술, 외국어의 한글 음역에 대한 것을 평가요소로 선택하였다.

4.2 측정방법

품질에 대한 지표와 평가요소를 설정하는 것도 어려운 일이지만, 각 요소를 측정하여 수량화하는 것은 더욱 어려운 작업이다. 데이터베이스의 품질을 평가하는 데는, 체계적으로 레코드를 표집하여 내용을 분석함으로써 평가요소에 대한 측정치를 실증적으로 제시하는 방법, 평가자들이 데이터베이스 안내자료를 조사하여 진술된 사항을 점검하는 방법, 그리고 탐색을 수행해 본 실패를 가지고 질에 대한 견해를 나타내는 방법 등 세 가지 방법이 있다. 본 모형에서는 세 가지 방법을 모두 적용하였다.

평가기준의 모형을 제시하거나 실제로 평가를 수행하는 연구에서 상당량의 연구가, 특히 WWW데이터베이스의 경우에는, 이용자 견지에서 데이터베이스의 성공정도 혹은 데이터베이스에 대한 만족정도를 평가하는 것이었다. 물론 이용자기반 평가도 상품과 서비스를 구축하거나 발전시키기 위한 기획과정의 한부분이다. 그러나 내용에 대한 보다 객관적인 평가는 일련의 표준적 테스트에 기초해야 하며, 피상적 평가가 되지 않기 위하여는 전문적으로 알고 이용할 수 있는 사람들에 의해 수행되어야 한다. 평가요소별로 적용된 실증적인 방법을 설명하면 다음과 같다.

포괄성은 오랫동안 문헌정보학에서 측정과 평가가 이루어져 왔다(Lancaster 1999). 포괄성은 특정주제의 정보가 해당 데이터베이스에 수록되어 있는 정도를 백분율(%)로 나타낸다. 서지나 본문 데이터베이스의 경우에, 학술지를 단위로 분석하는 방법과 논문을 단위로 분석하는 방법이 있는데, 논문을 단위로 분석하는 것이 더욱 정확하다. 포괄성의 측정방법은 서지대

(표 2) 지표별 평가요소와 측정방법

지 표	평가요소	Washington
1.포괄성	1. 포괄성 2. 레코드 수의 적절성 3 정보원의 적절성 4. 정보자료 유형의 다양성 5. 소급정도	서지대조법, 교차탐색(%) 5단계 척도 5단계 척도 5단계 척도 5단계 척도
2.최신성	1. 정보와 레코드 생산의 시차 2. 갱신빈도의 적절성	표집분석(일) 5단계 척도
3. 완전성	1 완전한 레코드의 비율 2. 구성필드의 적절성 3. 탐색필드의 적절성 4 색인어 통제 5. 통제어휘의 유지보완	표집분석(%) 5단계 척도 3단계 척도 2단계 척도 3단계 척도
4. 정확성	1 정확한 레코드의 비율	표집분석(%)
5.일관성	1 출판사항 기술 2 외국인 저자명 기술 3 약어의 기술 4. 외국어의 한글음역	5단계 척도 5단계 척도 5단계 척도 5단계 척도

조법, 유사 데이터베이스 교차탐색법, 인용문헌 대조법 등 여러 가지 방법이 있으나, 인용문헌 대조는 문헌의 선택에 문제가 있으므로, 서지대조법이나 유사 데이터베이스 교차탐색 방법이 적절하다. 서지대조법은 해당 주제에 관한 서지를 발견하여 서지에 있는 문헌이 평가할 데이터베이스에 수록되어 있는지의 여부를 조사하는 것이다. 우리나라의 경우 서지를 발견하기가 용이치 못하므로, 유사 데이터베이스를 각각 탐색한 후 탐색결과를 통합하여 일종의 서지를 만든다. 각 데이터베이스에 포함여부를 조사하여 포괄성을 측정하는 교차탐색 방법이 더욱 효과적일 것이다.

최신성, 완전성, 정확성을 측정하기 위하여는

조사할 레코드의 표집이 필수적이다. 통계적으로 유의한 결과를 얻으려면 체계적인 무선표집이 필요하며, 조사할 레코드의 수도 결정해야 한다. 표본의 크기에 대한 공식이 있지만, 모집단의 성질과 연구목적 등에 대한 고려는 매우 중요하며, 절대적인 기준은 존재하지 않는다.

우리나라의 데이터베이스는 수록기간이 매우 짧고(5년이하의 데이터베이스가 40.6%) 대규모의 데이터베이스는 드물다(한국데이터베이스진흥센터, 1999). 또한 유한 모집단에서 어떤 속성을 알고자할 때(예, 정확/부정확, 완전/불완전 등등) 그다지 큰 표본이 필요하지 않다. 통계표에 의하면 모집단이 1,000,000 그리고 10,000 일 때 거의 비슷한 수의 크기가(384,

370) 제시되어 있다(Krejcie and Morgan 1970). 표본의 수가 정해지면, 전체레코드 수를 표본의 수로 나누어 매 n(몇)번째의 레코드를 선택하여 분석하면 된다.

최신성은 각 표본 레코드에 대하여 정보가 생산된 일자(예, 출판일자) 해당 문헌이 데이터베이스 적재된 일자 간의 시차를 구하여 평균을 산출한다. 완전성과 정확성은 표본레코드 전체 중에서 완전한 레코드와 정확한 레코드의 비율을 산출한다.

위와 같은 실증적 측정을 적용할 수 없는 나머지 평가요소들은 데이터베이스 안내자료의 조사 및 점검, 그리고 샘플탐색을 수행하여 각각 내용에 따라 5단계, 3단계, 2단계 척도로 측정한다.

본 모형에서 설정한 5개 평가지표에 관한 평가요소와 측정방법을 정리하면 <표 2>와 같다.

4.3 평가척도의 개발

측정의 결과는 수치로 나타나고 평가요소 전체가 점수화되어 가능한 총점이 통상적인 1~100사이의 값으로 표현될 수 있어야 할 것이다. 이러한 방법은 아주 세밀한 구별을 가능케 하며, 특히 유사한 데이터베이스군을 비교할 때 효과적이다.

이 접근방법에 대한 한 가지 문제점은 지표별, 그리고 지표내 요소별 점수의 배정에 있다. 물론 제시된 지표와 평가요소들이 모두 동등하게 중요한 것이 아님은 자명하다. 그러나 여러 가지 지표에 배정할 상대적 가중치에 대하여 관계자들은 각각 다른 견해를 가질 수 있다.

선행연구에 나타나있는 연구자들의 지표별 중요도에 대하여 살펴보면 다음과 같은 사항을

찾아볼 수 있다.

Wilson은 전문가들을 대상으로 하여 SCOUG에 나타나 있는 데이터베이스의 품질지표를 선택하여 중요도에 따라 서열을 매기도록 요청하였는데, 그 결과를 보면 포괄성, 접근가능성, 최신성, 일관성, 정확성의 순서로 되어 있다(Rittburger 1997, 26). Armstrong은 SCOUG의 10가지 지표에 대하여 중요성과 측정가능성의 정도를 2차원 행렬로 나타냈는데, 그 순서는 정확성, 일관성, 최신성, 포괄성의 순서로 되어 있다(Armstrong 1995, 224). Harry와 Oppenheim(1993, 348-349)은 데이터베이스 평가를 위한 점수화를 시도함에 있어 최신성과 완전성에는 동일한 점수를 포괄성에는 보다 높은 점수를 배정하였다. Lancaster는 포괄성, 검색성, 최신성 등을 논의하면서 최신성의 중요도를 다른 지표보다 낮게 간주하였다(Lancaster 1999, 205). 이국희는 학술 데이터베이스의 경우 정확성을 가장 중요하게 보고, 그 다음이 완전성, 그리고 현행성과 일관성의 중요도를 동일하게 취급하였다(이국희 1998, 636).

<표 3>은 평가 지표별 가중치에 대한 연구자들의 견해를 정리한 것이다.

<표 3> 선행연구에 나타난 지표별 가중치

연구자	중요도(가중치)
Wilson	포괄성>최신성>일관성>정확성
Armstrong	정확성>일관성>최신성>포괄성
Harry & Oppenheim	포괄성>최신성 = 완전성 (4 1 1)
이국희	정확성>완전성>최신성 = 일관성 (4 3 2 2)
Lancaster	포괄성>최신성

측정가능성을 강조한 Armstrong의 연구를 제외하고는 포괄성이 가장 중요한 요소로 대두되었다. 포괄성은 데이터베이스가 다루는 주제에 대한 정보의 포함여부를 측정하는 것으로 내용지표에서 그 중요성은 의심할 여지가 없다. 정확성과 최신성은 업무의 효율과 관계가 있으며, 특히 정확성은 관련기법의 발달로 시스템 지원에 투자를 한다면 일정 수준을 달성하기가 비교적 쉬운 지표이다. 본 연구에서는 포괄성, 완전성, 일관성, 최신성, 정확성의 순서로 중요도를 설정하고 각각 6:5:4:3:2의 가중치를 부여하였다.

평가지표 내의 평가요소에 대하여는 내용자체를 보다 심도있게 측정하는 실증적 데이터에 보다 높은 가중치를 주었다. 해당 측정치에 대하여는 그 수준을 가늠할 준거가 필요하다. 본 연구에서는 공개된 자료와 선행연구에 나타난 결과를 토대로 척도를 마련하였다. 가능한 한 평균값을 중심으로 최대치와 최소치를 범위로 하여 5단계 척도를 적용하였다.

포괄성 측정(%)의 제한점은 바람직한 수준에 대한 절대값이 없다는 점이다. 특정 주제의 정보를 모두 수록할 수 있는 데이터베이스란 존재할 수 없다. 선행연구에 의하면 우리나라의 경우 포괄성은 33%~63%의 범위로 나타난 바 있으며(장혜란 1994), Sievert와 Sievert의 연구에서는(1991) 54%로 나타났다. Journal포괄성을 분석한 LaBorie등의 연구에서는(1985) 32%~79%, Jaguszewski와 Kempf의 연구에서는(1995)16%~93%, Giral과 Taylor의 연구에서는(1993) 72%로 나타났다. 본 모형에서는 포괄성의 범위를 30%~60%로 설정하였다.

최신성을 평가하기 위한 시차에 관한 사항은 대체로 데이터베이스제작자의 정책에 영향을

받는다. 권위있는 데이터베이스의 시차분석 결과를 보면 EMBASE의 경우는 문헌을 입수한 시점부터 데이터베이스에 적재될 때까지 시차가 3주로 나타나 있으며(Brigg and Growlesmith 1995), ABI-INFORM의 경우에는 문헌의 입수부터 데이터베이스에 적재될 때까지의 시간을 정책적으로 4주내로 한정하고 있다(Garman 1988). 또한 RNdexTop100 데이터베이스에서는 색인 작성에 소요되는 시간을 4주로 감축시킬 것을 제안하고 있다. 최신성은 제작자에게 문헌이 입수된 시점이 아니라, 문헌 생산 시점으로부터 데이터베이스에 적재될 때까지의 경과시간을 측정한다. 따라서 본 모형에서는 시차의 최소치를 5주로 하고 최대치를 9주로 범위를 35일~63일로 설정하였다. 텍스트데이터베이스 중에서 디렉토리 유형에는 시차를 측정하지 않고 갱신빈도의 적절성을 가지고 최신성을 평가한다.

완전성 평가에서는 완전한 레코드의 비율이 중요한 요소인데, 이것을 위하여 제시된 명백한 척도를 발견할 수 없었다. 90%~100%범위로 설정하였다. 색인은 서지 및 전문 데이터베이스의 품질에 중요한 영향을 미친다. 국내 데이터베이스의 색인을 조사한 연구결과를 보면 27개의 주요 데이터베이스중 8개만이 통제어휘를 사용하는 것으로 나타나 있다(정영미 1996). 따라서 이 평가요소는 색인통제의 존재와 유지의 적절성으로 측정하였다. 디렉토리 유형의 데이터베이스에는 통제어휘와 그 유지에 관한 요소는 적용하지 않는다.

정확성에 관한 외국의 연구에서는 대체로 저차명 기입의 오류(4.5%), 서지사항 기술상의 오류가 9.4%(Pao 1989)로 나타나 있다. Bourne은 11개의 데이터베이스를 대상으로 철

자오류를 조사하였는데, 색인어의 철자오류는 0.5%~ 23%에 이르며, 1000레코드당 부정확한 서지사항을 가진것은 0.125~6.25에까지 이른다(Bourne 1977). 최고 품질의 경우 철자오류는 0.2%(1000단어 당 2개, O' Neill & Vizine Goetz 1988)로 나타났다. 데이터베이스 크기가 클수록 오류의 수는 많아지므로, 전체 혹은 부분에 대한 오류의 측정은 의미가 없다. 본 연구에서는 표집된 레코드를 단위로 하여 오

류의 유무를 측정하며 정확성의 범위는 96%~100%로 삼았다.

일관성에 관한 사항은 표집에 대하여 일반적으로 적용할 수가 없다. 따라서 데이터베이스제작자의 표준화에 대한 진술내용을 샘플탐색을 통하여 검증하고 5단계 척도로 표시하도록 만들었다.

〈표 4〉는 지표별 평가요소에 대한 평가척도를 정리한 것이다.

〈표 4〉 지표별 평가척도

평가지표	평가요소	점수	척 도				
			1	2	3	4	5
포괄성(30)	포괄성(%)	10	30이하	30-39	40-49	50-59	60이상
	레코드 수	5	매우부적절	약간부적절	보통	약간적절	매우적절
	정보원	5	매우부적절	약간부적절	보통	약간적절	매우적절
	정보자료유형의 망라성	5	매우부적절	약간부적절	보통	약간적절	매우적절
	소급정도	5	매우부적절	약간부적절	보통	약간적절	매우적절
최신성(15)	시차(일)	10	57이상	50-56	43-49	36-42	35이하
	갱신빈도	5	매우부적절	약간부적절	보통	약간적절	매우적절
완정성(25)	완전한 레코드 비율(%)	10	91-92	93-94	95-96	97-98	99-100
	구성필드	5	매우부적절	약간부적절	보통	약간적절	매우적절
	탐색필드적절성	3	부적절		보통		매우적절
	통제어휘	4				존재	
정확성(10)	통제어휘갱신	3	부적절	보통	매우적절		
일관성(20)	정확률(%)	10	96미만	96-96.9	97-97.9	98-98.9	99이상
	출판사항	5	매우부적절	약간부적절	보통	약간적절	매우적절
	외국인저자명	5	매우부적절	약간부적절	보통	약간적절	매우적절
	약어표준화	5	매우부적절	약간부적절	보통	약간적절	매우적절
	외국어의 한글 음역	5	매우부적절	약간부적절	보통	약간적절	매우적절

5 결론 및 제언

이상과 같이 데이터베이스 품질에 관하여 선행연구 결과를 기초로 하여 내용을 평가하기 위한 지표를 설정하고, 지표별 측정요소와 측정방법을 확립한 후, 척도를 제시함으로써 구체적인 평가 모형을 완성하였다.

본 연구에서 제시된 모형을 이용하여 각 데이터베이스의 품질을 측정할 수 있으며, 평가결과는 해당 데이터베이스가 특정 지표 혹은 특정 지표내 요소에서 어느 정도 성공적인지를 판단할 수 있게 한다. 이 모형은 온라인과 CD-ROM 형태의 텍스트데이터베이스에 모두 적용가능하다. 모형을 구축함에 있어서, 문헌정보학 분야에서 발전되어 온 표준적인 측정방법을 가능한 적용코자 노력하였으며, 실증적인 측정치에 보다 많은 가중치를 두었다. 그러나 제시된 모형은 완전한 것이 아니며 시험적용을 통하여 제한점에 대한 피드백이 제공되고 수정이 이루어질 수 있을 것이다.

주지하는 바와 같이 우리는 모두 평가의 중요

성을 인식하고 있으며, 또한 사회전반에서 부각되고 있는 평가요구를 더 이상 외면할 수 없다. 평가는 하나의 연구이며 평가의 수행은 시간과 노력, 즉 비용이 드는 어려운 작업이다. 그러나 일단 시작이 되면, 차후에는 신규 데이터베이스를 제외한 기존 데이터베이스에 대한 평가는 갱신부분을 대상으로 평가함으로써 보다 용이한 작업이 될 것이다.

정기적으로 주요 데이터베이스들이 평가되면, 그 측정치의 평균과 표준편차를 가지고 품질의 등급을 규정할 수 있으며, 평가 결과는 공표되고, 궁극적으로는 권위있는 기관에 의하여 인정될 수 있을 것이다 이와같은 데이터베이스 평가의 제도화는 탐색자가 데이터베이스를 올바르게 선택할 수 있는 정보를 제공함으로써, 사회내에서 질적으로 낮은 데이터베이스가 도태되고 양질의 데이터베이스가 활성화됨과 동시에, 데이터베이스제작자가 고품질의 데이터베이스를 개발 유지할 수 있도록 장려함으로써, 생산적인 정보사회의 도래를 촉진시킬 수 있을 것이다.

참고 문헌

- 김지훈, 1996. 정보서비스의 품질평가에 관한 고찰. 『도서관학 논집』, 25: 441-474.
- 유사라, 1999. 데이터베이스 정보 품질 평가의 메타분석. 『정보관리학회지』, 16(1): 157-174.
- 유사라, 2000. Web 정보서비스 평가를 위한 기준 측정지표 분석 I. 『한국문헌정보학회지』, 34(3): 133-155.
- 유혜영, 1997. 『국내 제작 데이터베이스의 평가에 관한 연구』, 석사학위논문, 서울여자대학교 대학원, 문헌정보학과.
- 이국희, 1998. 데이터베이스 품질 평가에 관한 연구. [cited 2000. 9. 15.]. <<http://www.dpc.or.kr/report/quality.html>>
- 이제환, 1997. 과학기술분야 서지 DB의 품질관리 및 평가 방안: KORDIC의 KRISTAL DB를 중심으로. 『한국문헌정보학회지』, 31(3): 109-134.
- 이제환, 1998. 분산체제로 구축된 통합DB의 품질 관리에 관한 연구. 『한국문헌정보학회

- 지』, 32(3): 179-206. 한국데이터베이스진흥센터. 알기쉬운 한국의 데이터베이스 목록. 1999. pp. 6-37. 서울
- 장혜란. 1994. 정기간행물 기사 색인 서비스의 포괄성 측정에 관한 연구. 『정보관리학회지』, 11(1): 17-29
- 정영미. 1996. 국내문자 전문 데이터베이스의 색인에 관한 연구. 『정보관리학회지』, 13(1): 19-45.
- Armstrong, C. J. 1995. "The Eye of the Beholder." In: Basch R. ed. *Electronic Information Delivery*. Vermont: Gower. 221-244.
- Armstrong, C. J. 1996. "Questions of Quality." *The New Review of Information Networking*, 249-253.
- Azgalov, E. G. 1969. "A Framework for Description and Classification of Printed Subject Indexes." *Libri*, 19: 275-291.
- Basch, R. 1990. "Measuring the Quality of the Data: Report on the Fourth Annual SCOUG Retreat." *Database Searcher*, 6(8):18-23.
- Basch, R. ed. 1995. *Electronic Information Delivery*. Vermont: Gower.
- Bearman, T. C. and Kunberger, W. A. 1977. *A Study of Coverage, Overlap among Fourteen Major Science and Technology Abstracting and Indexing Services*. Philadelphia: National Federation of Abstracting and Indexing Services.
- Bourne, C. P. 1977. "Frequency and Impact of Spelling Errors in Bibliographic Databases." *Information Processing & Management*, 13(1): 1-12.
- Brennan, P. B. et al. 1999 "What Does Electronic Full-text Really Mean? A Comparison of Database Vendors and What They Deliver." *Reference Services Review*, 27(2): 113-126.
- Briggs, K. and Growlesmith, I. 1995 "Embase-The Excepta Medica Database: Quick and Comprehensive Drug Information." *Publishing Research Quarterly*, 11(3): 51-60
- CIQM. 1996. "A Description of Database Labels." [cited Oct. 15, 2000] <<http://www.la-hq.org.uk/liaison/ciqm>>.
- Cluley, H. J. 1969. "Analytical Abstracts: User Reaction Study." *Proceedings of the Society for Analytical Chemistry*, 5: 217-221.
- CyberHound's Guide to Internet Databases*. [database online]. Detroit: Gale Research, 1994-, Available from Dialog File No.230, Palo Alto, Calif.
- Dong, X. and Su, L. T. 1997. "Search Engines on the World Wide Web and Information Retrieval from the Internet: a Review and Evaluation." *Online & CD-ROM Review*, 21(2): 67-82.
- Drage, J. E. 1969. "User Preferences in Technical Indexes." *The Indexer*(6): 151-155.

- Foulds, M. S. and Foulds, L. R. 1990. "CD - ROM Disk Selection and Evaluation " *Reference Services Review*, 18(2): 27-38.
- Garman, N. 1988. "An Inside Look at an Online Database." *Database* (Apr. 1988): 50-56.
- Giral, A. and Taylor, A. G. 1993. "Indexing Overlap and Consistency between the Avery Index to Architectural Periodicals and the Architectural Periodical Index." *Library Resources and Technical Services*, 37(1): 19-44.
- Hall, A. M. 1972 *User Preferences in Printed Indexes*. London: Institute of Electrical Engineers.
- Harris, G. 1990. "The User Interface to Online Information." *Proceedings of International Online Information*, (14): 319-324.
- Harris, R. 1997. "Evaluating Internet Research Sources." [online] [cited Oct, 10, 2000]. (<http://www.sccu.edu/faculty/R_Harris/evalu8it.html>).
- Harry, V. and Oppenheim, C. 1993 "Evaluations of Electronic Databases, part I : Criteria for Testing CD ROM Products." *Online & CD ROM Review*, 17(4): 211-222.
- Harry, V. and Oppenheim, C. 1993. "Evaluation of Electronic Databases, part II : Testing CD ROM Products." *Online & CD ROM Review*, 17(6): 339-350.
- Herther, N. 1988. "How to Evaluate Reference Materials on CD-ROM," *Online*, 12(2): 106-108.
- Jacso, P. 1992. "What is an Update? Currency Test Search of Databases." *Database* 15: 28-33.
- Jacso, P. 1997. "Content Evaluation of Databases." *Annual Review of Information Science and Technology*, 32 231-267
- Jaguszewski, J. A. and Kempf, J. L. 1995, "Four Current Awareness Databases: Coverage and Currency Compared." *Database*, 18(1): 34-44.
- Juntunen, R. et al. 1991. *Quality Requirements for Databases-project for Evaluating Finnish Databases*. London: Online Information.
- Kirkwood, H. P. 1998. "Beyond Evaluation: A Model for Cooperative Evaluation of Internet Resources." *Online*, 22(4): 66-72.
- Krejcie, R. V. and Morgan, D. W. 1970. "Determining Sample Size for Research Activities." *Educational and Psychological Measurement*, 30: 608.
- 재인용: Powell, R. R. 홍현진 역. 『문헌정보학의 연구방법론』. 서울: 구미무역, 1996, pp.130-133.
- LaBorie, T. et al. 1985. "Library and Information Science Abstracting and Indexing Services: Coverage, Overlap and Context." *LISR*, 7: 185-195.
- Lancaster, F. W. 장혜란 역. 1999. 『색인초록의

- 이론과 실제』. 서울: 구미무역.
- Martyn, J. and Slater, M. 1964. "Tests on Abstracts Journals." *Journal of Documentation*, 20: 212-235.
- Mintz, A. P. 1990. "Quality Control and The Zen of Database Production." *Online*, 14(6): 15-23.
- Nicole, A. 1998, "Bibliography on Evaluating Internet Resources." [cited Sept. 10, 2000]. <http://refserver.lib.vt.edu/libinst/crit/THINK.HTM>. Virginia Polytechnic Institute and State University.
- O' Neill, E. T. and Vizine-Goetz, D. 1988. "Quality Control in Online Databases." *Annual Review of Information Science and Technology*, (23): 125-156.
- Oppenheim, C. et al. 2000. "The Evaluation of WWW Search Engines." *Journal of Documentation*, 56(2): 190-211.
- Pao, M. L. 1989. "Importance of Quality Data for Bibliometric Research." *Proceeding of the National Online Meeting*, (10): 321-327.
- Rittberger, M. and Rittberger, W. 1997. "Measuring Quality in the Production of Databases." *Journal of Information Science*, 23(1): 25-37.
- Sievert, M. C. and Sievert, D. E. 1991. "Online Searching in Philosophy : A Comparison of Philosophers Index and Francis." *Online Review*, 15(2): 63-76.
- Smith, A. G 1997 "Testing the Surf: Criteria for Evaluating Internet Information Resources." *The Public Access Computer Systems Review* [online], 8(3)
- Vercellesi, L. 1993. "Quality Assurance in Drug Information Centres." *Online & CD ROM Review*, 17(5): 279-282.
- Wang, R. Y., Storey, V. C., and Firth, C. P. 1995. "A Framework for Analysis of Data Quality Research." *IEEE Transactions on Knowledge and Data Engineering*, 7(4): 623-640.
- Wilson, T. 1994. *EQUIP: A European Survey of Quality Criterion for the Evaluation of Databases*
- Zink, S. D. 1991. "Towards More Critical Reviewing and Analysis of CD-ROM User Software Interfaces." *CD-ROM Professional*, 4(1): 16-22.