

데이터베이스 정보 품질 평가의 메타분석*

Meta-Analysis of Domestic ST Database Evaluation

유사라(Sa-Rah Yoo)**

목 차

1 서론	4 결과 분석
2 기존연구	5 결론
3 평가 실험탐색	

초 록

본 연구는 국내에서 제작된 주요 과학기술 데이터베이스 3개를 선정하고 각 데이터베이스의 검색결과를 조사하여 데이터베이스의 품질을 종합적으로 평가한 메타 연구이다. 종합적 평가로부터 지적된 문제점들을 나열하고 해결방안과 정책적인 제언을 모색하고자 하는 목적으로 진행되었다. 연구의 주요 결과는 첫째, 기준측정에서 상기한 연구대상들은 적지 않은 오류와 미비점을 보이고 있으며 둘째, 일반적으로 데이터베이스자체에 대한 공식적인 설명이나 안내가 누락되어 있고, 셋째, 검색되는 색인어나 주제어의 용어통계가 완벽하게 이루어지지 않고 있다. 넷째 최종 이용자 검색을 지원하고 있는 색인어 브라우징이나 주제별 상호참조와 같은 서비스 지원이 제대로 이루어지고 있지 않으며, 마지막으로 검색시 발생할 수 있는 오류와 그에 대한 이용자의 오류처리안내나 설명이 검색과정에서 제공되지 않거나 실질적인 도움을 주지 못하고 있는 것으로 나타났다.

ABSTRACT

This investigation used meta-analysis to examine the quality of data of domestic ST(Science & Technical) databases. Analyses measured the quality of data itself, determined whether these databases have any critical errors in terms of accuracy, completeness, and consistency by conducting the data retrieval experiments. The purpose of this study was to diagnose the science & technical databases and to find out the problems of them and also to provide possible suggestions for their improvements. The researcher concluded the results by mentioning the three research questions.

키워드: 과학기술 데이터베이스, 품질 평가, 완전성, 정확성, 일관성, Meta-analysis, Accuracy, Completeness, Consistency, Domestic Science & Technical Databases.

* 한국학술진흥재단 1997년도 자유 공모과제로 지원된 연구임.

** 서울여자대학교 경상정보학부 문헌정보학과 교수

■ 논문접수일 : 1999년 3월 4일

1 서론

일반적으로 기존의 연구되어진 정보 시스템의 이론적인 효과성과 효율성 측정은 정보검색시스템 구성소 각각의 평가측면에서 요약될 수 있다. 정보시스템 구성소는 기존 연구들에 의하여 다양하게 나열될 수 있으나 크게 정보자원인 데이터베이스 요소, 정보검색엔진(기법)요소, 이용자 인터페이스 요소, 이용자 만족도 요소, 정보접근·전달요소로 볼 수 있다. 정보시스템의 정보자원인 데이터베이스 요소는 여러 측면에서 다양한 모양과 형태로 생산되고 있으며, 직·간접적으로 온라인 정보시스템의 제기능에 큰 영향을 주고 있다. 비록, 검색엔진이 훌륭하고 이용자 인터페이스가 최선의 정보기술로 제공되고 있으며 초고속 네트워크에 의한 정보전송이 가능해도 실제 검색환경에서 이용자에게 주어지는 데이터베이스 정보의 질적 수준이 미흡하거나 접근 메커니즘이 제기능을 못하면 정보시스템의 효과·효율성과 이용자 만족도는 기대될 수 없다. 이미 이용자가 경험한 전자정보 데이터의 오류와 질적 수준에 대한 불신을 제거되지 못한 상태에서, 더군다나 검색환경에서는 직접적인 변화를 쉽게 볼 수 없는 디지털 정보에 대한 이용자의 기대는 막연한 것이 아닐 수 없다. 따라서 정보전문가들은 현재 우리가 갖추거나 가공하려고 막대한 비용을 투입하는 데이터베이스가 이용자에게 정보로 받아들여지고 있는지, 그리고 신뢰할 수 있는 데이터베이스로 인식되고 있는지를 확인해야 하고 그에 대한 정확한 대답을 얻어야 하는 작업을 해야 한다.

1.1 연구 문제

국내에서 제작되고 유지되며 과학기술 연구집

단에게 서비스되고 있는 중요한 과학기술 데이터베이스를 대상으로 그들이 갖고 있는 정보 자체의 품질을 평가하는 것을 목적으로 한다.

1. 국내에서 제작되고 있는 과학기술연구정보의 데이터베이스 품질은 어떠한가?

2. 데이터베이스 정보자체의 품질측정에서 지적된 문제들은 무엇이며 이를 해결하기 위해 적용될 수 있는 방안들은 어떠한 것들이 있는가?

3. 국내 과학기술연구 정보의 데이터베이스 품질개선을 위해 제시될 수 있는 최우선적인 정책이나 대안은 무엇인가?

본 연구에 사용된 연구방법은 메타연구의 일종으로 선정된 연구대상 각각에 대한 자료조사와 검색이 적용되었으며 연구대상에 관한 최근의 기존연구들을 포함하여 데이터품질에 대해 종합적인 결론을 유도하는 방법으로 진행되었다.

1.2 연구범위

연구범위는 국내에서 제작한 과학기술정보 데이터베이스의 정보자체에 대한 품질 평가이며 이것은 평가기준인 완전성, 정확성, 일관성의 3측면의 실험검색으로 측정하고 분석했다. 평가의 기준은 Reva (1990)의 데이터베이스 정보평가의 방법으로 ① 데이터의 정확성과 부정확성, ② 데이터의 일관성과 비일관성, ③ 데이터의 완전성과 불완전성을 측정했다.

정확성이란 레코드에 저장된 데이터의 오류가 없는가를 의미하는 것으로, 데이터 자체의 잘못된 표기나 원시 자료(raw data) 내용에 대한 부정확성의 반대 개념이다. 정확성은 원정보의 확실성에 준거하며 탐색자들은 검색결과에 정보에 대하여 신뢰할 수 있어야 한다. 따라서 데이터베이스의 부정확성이란 데이터베이스 필드구조상의

설계방법이나 색인방식등의 이유로 인해 오히려 찾고자하는 적합 정보를 얻지 못하게 되는 가능성을 의미한다.

일관성은 데이터베이스내 둘 이상의 데이터 값이 서로 상충되지 않고 일관된 상태를 말한다. 동일한 사실을 다양한 용어로 표현할 수밖에 없는 경우에 대한 전거통제나 통제어휘의 일관성 정도를 의미한다. 일관성은 데이터베이스 이용자들이 검색에서 얻을 수 있는 정보와 얻지 못할 정보에 대한 사전의 범위를 미리 예측할 수 있게 하는 중요한 도구의 역할을 한다. 데이터베이스내의 모든 레코드의 정보가 일관성이 있다면 소수의 필드만으로도 원하는 정보를 쉽게 얻을 수 있다.

완전성은 데이터베이스에 수록된 정보의 포괄성 정도나 가공된 데이터가 원문에 담긴 정보를 완전하게 담고 있는지를 의미하는 것으로 크게 레코드의 범위, 구조 그리고 정보 접근성의 정도를 포함하는 개념이다.

1.3 연구대상

국내 과학기술연구 데이터베이스 정보를 제작하고 있는 기관은 대학을 비롯해 정부산하 연구기관, 특정주제 연구소등이 다양한 성격과 각각의 유일한 기능을 갖고 있다. 그러한 기관중 본연구에서 선정된 연구 대상은 그 성격에 따라 세가지로 구분할 수 있다.

첫째 연구대상은 과학기술대학으로 교육기관의 성격을 갖는 교육기관인 한국과학기술원의 과학도서관 (KAIST: Korea Advanced Institute of Science and Technology - Science Library)으로 이곳은 1971년 한국과학원(KAIS)도서실로 시작하여 한국과학기술연구원 도서실과의 통합분리 과정을 거쳐 1990년이래 과학기술대학(KIT)

과 통합운영되고 있다.

두 번째 대상은 연구기관의 성격을 갖는 기관으로 KORDIC(Korea R&D Information Center)을 선정하였으며 이는 과학기술처 산하 한국과학기술원 부설 연구기관인 연구개발정보센터 (KORDIC: Korea R&D Information Center)에서 제작한 과학기술 종합 데이터베이스인 KRISTAL 정보 서비스(Korea Research Information of Science & Technology Access Line)를 제공하고 있다. KRISTAL은 연구개발정보센터가 과학기술처 산하 정부 출연 연구기관과 함께 과학기술정보를 수집, 분석, 가공, 구축하여 서비스하는 체제이며 정보 검색 시스템 KRISTAL-II에 의해 인터넷, 연구전산망 (KREONet), 공중정보통신망을 통해서 제공되고 있다. KRISTAL에서 제공하는 데이터베이스는 크게 7개 분야로 나눌 수 있다. 과학기술 문헌/소재정보, Factual 정보, 과학기술 공용 정보, 과학기술 인력 정보, 국가 연구과제 정보, 국내 학회 학술 정보, 해외 과학기술 정보 등으로 구분된다.

세 번째는 과학기술관련 산업계에서 산업정보를 지원하고 있는 산업기술정보원 (KINITI: Korea Institute of Industry and Technology Information)으로 이곳은 70년대 후반부터 시작되었다고 할 수 있는 초기 정보화시대부터 과학기술 관련 데이터베이스를 제작하고 있는 기관이다. 산업기술정보원의 국내제작 데이터베이스는 문헌 데이터베이스로, 산업재산권 데이터베이스, 기술시장 데이터베이스, 총람 데이터베이스, 경제 데이터베이스 등이 있다.

1.4 연구의 제한점

본 연구는 다음과 같은 몇 가지 제한점이 있음

을 밝힌다. 첫째, 데이터베이스의 품질을 궁극적으로 판단하고 결정하는 사람은 이용자인데 본 연구에서 제안한 평가 기준과 그 방법 및 도구들은 무엇보다 이용자 입장에서 평가하는 것을 전제하고 있다. 그러나 이용자입장에서는 가장 중요한 요소의 하나인 현행성에 대한 측정이 연구대상의 대부분이 이에 대한 정확한 정보를 갖추고 있지 않은 관계로 제외되었다.

둘째, 연구 대상으로 선정된 데이터베이스들은 실험검색이 가능한 데이터베이스로서 각 기관에서 제공하는 모든 데이터베이스를 모두 포함하고 있지 않다. 예를 들면 KORDIC의 경우에 문헌데이터베이스에 해당되는 과학기술 도서정보(UNION), 과학기술 문헌정보(SATURN), 정보과학회 논문지(KISS), 해외과학기술 동향지(TECTREND)등으로 대상을 제한하였다. 이 외에 사실정보, 과학기자재정보, 인력정보 등의 데이터베이스는 시간의 흐름에 따라 실제 값(value)이 지속적으로 변하며 원정보의 확실성에 따라 데이터베이스 품질 평가가 달라지므로 본 연구 대상에서는 제외시켰다. KINITI는 총 19개의 데이터베이스 중에서 문헌 데이터베이스인 5개 데이터베이스인 BIST, DIGS, DIMD, INFO, KSMA로 검색을 제한하였으며 각 데이터베이스의 특성에 따라 검색용어나 검색방법에 제한을 받고 있어 실험용어는 일반적인 오류발생을 예상할 수 있는 것으로 선정되었다.

셋째, 데이터베이스 품질 평가에서 가장 큰 난점은 원시자료(raw data) 자체를 검증하고 확인하기가 어렵다는 점이다. 데이터베이스내 대규모 데이터 실제 값을 정확히 파악하는 것 자체가 시간과 비용의 문제로 인해 단일 연구로는 현실적으로 불가능하기 때문에 원시 데이터의 정확성 확인 및 평가는 측정범위에서 제외시켰다.

2 기존연구

데이터베이스 검색 결과의 품질에 영향을 미치는 요인을 오네일(O'Neill)은 수록된 데이터 자체의 품질요소와 이용자의 검색기능요소의 두 가지로 구분하고 수동적인 품질관리 방법의 요소로서 교육(education), 검토(review), 이용자의 협조(user assistance), 능동적 관리 방법으로 자가 점검 데이터(Self-Checking Data)의 이용, 자동 데이터 검증(Automated Data Validation), 그리고 전거통제(Authority Control)를 제안하였다.

데이터베이스 품질에 대해 주요한 논의를 해온 National Online Meeting, Online/CD-ROM 90 그리고 International Online Information Meeting등 국제회의에서는 공식적으로 '데이터베이스 품질'을 1990년의 슬로건으로 선언하면서 전세계적으로 이에 대한 보다 신중한 관심을 유도했다. 구체적 평가도구로 우수하다고 평가되어 오고 많이 적용되고 있는 것들을 살펴보면 SCOUG(The Southern California Online User's Group), 핀란드의 피니쉬 프로젝트(The Finnish Society for Information Services), 그리고 Peter Jacso가 1992년 개발한 CD-ROM 평가법, 캐롤 테노퍼(Carol Tenopir)가 개발한 평가법, AT&T Bell 연구소에서 개발한 평가법등을 들 수 있다. 이 평가법들의 공통된 특징은 데이터베이스의 질적 측면을 단일 기준으로 평가하지 않고 다양한 기준으로 나누었다는 점이다.

우선 SCOUG는 서지 데이터베이스나, 본문(full-text), 디렉토리형의 데이터베이스 출력물이나 검색결과를 중심으로 평가방법을 개발하였으며 탐색자관점에서 데이터의 품질을 평가할 수 있는 요소를 제시하고 있다.

피니쉬 프로젝트 (The Finnish Society for

Information Services)에서 제시한 평가는 서지 데이터베이스에 한하여 적용하는 것으로 데이터베이스의 통신과 시스템간의 접속여부, 검색어와 검색을 지원하는 여러 가지 기술들, 데이터의 내용적 품질(content quality), 정보 검색시에 제공되는 도움말 기능, 검색비용(cost) 5가지 평가항목을 제안하고 있다.

CD-ROM 데이터베이스를 평가대상으로 자스코(Jasco)가 개발한 평가법은 4개의 평가기준인 소프트웨어, 데이터웨어, 데이터베이스, 그리고 하드웨어를 적용하고 있다:

1987년에 온라인 데이터베이스의 품질 관리(quality control)를 가장 주된 논제로 삼았던 테노피(Tenopir)는 가장 큰 영향을 주는 5 가지를 범위, 구조, 접근성, 정확성, 일관성으로 제시하고 있다.

AT&T Bell 연구소가 제시하고 있는 네가지 평가기준을 보면 정확성, 현행성, 완전성, 일관성이다.

국내에서는 한국데이터베이스진흥센터에서 제시한 평가기준은 크게 데이터베이스 데이터 품질 기준과 데이터베이스 서비스 품질 기준으로 나누고 다음과 같이 구분하고 있다.

가. 데이터베이스 데이터 품질 기준

- ① 정확성(accuracy)
- ② 완전성(completeness)
- ③ 현행성(currentness)
- ④ 일관성(consistency)

나. 데이터베이스 서비스 품질 기준

- ⑤ 검색성(searching)
- ⑥ 사용용이성(ease of use)
- ⑦ 사용자지원성(customer support)

평가기준에 대한 대부분의 연구에서 데이터 자

체의 품질에 공통적으로 적용하고 있는 몇가지 기준을 살펴보면 정확성, 일관성, 완전성, 그리고 현행성등으로 나타나고 있다.

3 평가 실험탐색

연구대상으로 선정된 데이터베이스들의 검색실험은 본 연구자와 대학원생 두명의 연구보조원에 의하여 각각 개별적인 연구과정으로 실시되었다. 연구대상에 대한 종합적인 매크로 평가는 1998년 5월부터 9월까지 완성되었다. KINITI의 경우는 97년 9월부터 97년 10월까지 기준에 실시되었던 자료조사와 검색결과를 분석하였고 KINITI의 일차검색은 유혜영(1997) 논문에서 본 연구자 지도하에 분석되었으며 98년 2월에 매크로 평가를 위한 재 검색을 실시하였다. KORDIC 자료는 김선형(1997) 논문에서 일차로 별도 검색, 분석되었고 KORDIC 데이터베이스의 기존자료 역시 97년 11월부터 12월까지 조사 검토된 후, 98년 4월에 재 검색이 실시되었다. 반면에 KAIST의 자료는 일차자료로서 실험검색에 의한 측정이 98년 5월부터 8월까지 실시되었다. 각 실험검색에 의한 검색결과는 개별 분석된 후, 평가기준에 따라 평가표로 종합되었다.

4 결과 분석

4.1 정확성 평가

정확성에 관한 첫번째 측정으로 외래어의 경우 표기방법에 따른 용어의 정확성을 실험한 결과, 동일한 의미의 외래어이지만 표기방법에 따라 검

색 결과는 다르다는 것을 보여 주었다. 결과에서 주목할 만한 것으로 <표 1> KAIST와 <표 2> KORDIC의 OR연산자 적용건수이다. SATURN과 TECTREND는 두 용어의 총합과 OR연산자를 적용한 문헌 총수가 큰 차이를 보이고 있으며 한 레코드내에 두가지 용어가 혼합 표기되고 있는 경우도 있다. 이것은 외래어가 여러 가지로

표기될 수 있기 때문에 한 레코드내에서 혼합 표기방식은 오류일 수밖에 없다. 완전한 검색을 위해서는 여러 가지 표기방법이 사용되고 있는 용어일지라도 이를 통제할 수 있어야 하며, 또한 어떠한 기입방식에 의해 입력된 용어라도 하나의 전거하에서 통제가 되고 이에 관한 모든 데이터는 검색되어야 한다. '네트웍'과 '네트워크' 용어

<표 1> KAIST 데이터베이스 정확성 평가표

DB명	검색용어	검색결과	or 연산자 적용 건수	외래어표기 오류여부
KAIST	윈도	6	180 (199)	초록필드상 문법오류 문장 외래어 철자표기오류
	윈도우	96	11건차이	
	윈도우즈	97		
	디지털	178	388 (419)	
	디지털	241	31건차이	

<표 2> KORDIC 데이터베이스 정확성 평가

DB명	검색용어	검색결과	or 연산자 적용 건수	외래어표기 오류여부
UNION	디지털	463	570	외래어 철자표기오류 'government' 'koenig'
	디지털	107		
	네트웍	47		
	네트워크	285		
SATURN	디지털	155	578 (691)	외래어 철자표기오류 'government' 서명란 pas와 psa의 혼용
	디지털	436	13건차이	
	네트웍	98	1503 (1513)	
	네트워크	1415	10건차이	
KISS	디지털	13	14	외국어철자 오류 초록필드 'semantics'
	디지털	1		
	네트웍	6	80 (81)	
	네트워크	75	1건차이	
TECTREND	디지털	13	552 (589)	
	디지털	552	37건차이	
	네트웍	116	947 (976)	
	네트워크	858	2건차이	

〈표 3〉 KINITI 데이터베이스의 정확성 평가표

DB명	검색용어	검색결과	or 연산자 적용 건수	외래어표기 오류여부
BIST	윈도	14	431 (438)	DT상 I문자표기오류, LA, 기사구분필드 데이 터 누락
	윈도우	351	7건차이	
	윈도우즈	73		
	데이터베이스	1122	1251 (1271)	
	데이터베이스	1150	21건차이	
INFO	윈도	2	18	초록필드에서 문법오류 문장
	윈도우	14		
	윈도우즈	2		
	데이터베이스	718	719 (723)	
	데이터베이스	5	4건차이	
DIGS	윈도	1	30 (31)	199년도 DB에서 1997 년도 자료검색
	윈도우	26	1건차이	
	윈도우즈	4		
	데이터베이스	156	375 (388)	
	데이터베이스	232	13건차이	
DIMD	윈도	0	45	
	윈도우	41		
	윈도우즈	4		
	데이터베이스	32	509 (512)	
	데이터베이스	480	3건차이	
KSMA	database	42		LA,DT필드상에 철자표 기오류
	databases	12		

로 실험한 결과에서도 역시 같은 오류가 발생하였다.

정확성의 두번째 측정방법은 탐색자가 색인어를 검토할 수 있도록 브라우징 기능이 제공되고 있는지에 관한 것으로 KORDIC의 경우에, UNION에서만 색인어의 브라우징이 가능하고 나머지 데이터베이스에는 그 서비스가 제공되지 않고 있다. 오류 데이터의 여부에 관해서는 TECTREND를 제외한 다른 데이터베이스에서 유형의 차이는 있었지만 오류로 입력된 데이터를

발견할 수 있었다.

KINITI의 경우, 'OR' 연산자를 사용한 각 용어별 검색문헌의 총합과 문헌 총수의 비교에서는 DIGS와 INFO가 결과상 차이를 보이지 않고 있다.

4.2 일관성 평가

일관성 평가를 위해 주제에 대한 일관성 있는 지원 방식인 상호 참조면에서의 일관성유지를 위

한 지원 여부에 대해 조사하였다. 실험 용어로 선정된 데이터베이스라는 용어는 'database', 'data base', 'DB' 등 세 가지 형태로 표기가 되고 있다. 일관성 평가의 두 번째 측정으로 대소문자의 구별과 띄어쓰기 인식의 여부에 대한 실험은 거의 모든 데이터베이스가 대소문자를 구별하지 않고 동일하게 취급한 것으로 나타났다. 이것은 탐색시에 전거통제나 통제어휘 등이 제대로 이행되지 않고 있으며 저자나 원시 자료에 의해

의도된 대소문자에 따른 의미상의 차이가 구별되지 못하며 다만 탐색자의 판단에 의해서만 가능하게 된다.

KORDIC의 경우, 모든 데이터베이스가 이 세 가지 표기방식을 혼합하여 기입하고 있었고 각 용어별로 검색된 결과도 다른 문서를 검색했다. 동일한 의미를 갖는 세 가지 형태의 용어를 전혀 다른 의미로 각각 인식하고 있기 때문에 색인어의 비 일관적인 표기방식은 검색 결과에 문제가

〈표 4〉 KAIST 데이터베이스 일관성 평가표

DB명	주제별상호참조	약어처리	대소문자 구별	띄어쓰기
KAIST	없음	DB 36	동일취급 ibm 436 IBM	database error data base 255 cd-rom 83 cd rom 4

〈표 5〉 KORDIC 데이터베이스 일관성 평가표

DB명	주제별상호참조	약어처리	대소문자 구별	띄어쓰기
UNION	없음	DB 검색 85건	동일취급 UNION union 1370	database 688 data base 799
SATURN	없음	DB 검색안됨 레코드상에는 약어나타남	동일취급 UNION union 6170	database 6097 data base 2037
KISS	없음	DB 검색 16건	동일취급 BASIC union 1370	database 136 data base 16
TECTREND	없음	DB 검색 77건	동일취급 UNION union 12	database 20 data base 1

〈표 6〉 KINITI 데이터베이스 일관성 평가표

DB명	주제별상호참조	약어처리	대소문자 구별	띄어쓰기
BIST	없음	DB 276건 (database 75)	IBM 88 ibm 327 OCLC 1 oclc 0	cd-rom 41 cd rom 34
INFO	없음	DB 42건 (database 76)		
DIGS	없음	DB 106건 (database 2070)	동일취급	cd-rom 44 cd rom 15
DIMD	없음	DB 15건 (database 73)	동일취급	cd-rom 25 cd rom 0
KSMA	없음	DB 51건 (database 64)	동일취급	cd-rom 1 cd rom 0

있음을 보이고 있다.

KINITI의 경우, 주제적인 측면에서 일관성을 부여하는 주제별 상호 참조가 있는지를 조사한 결과, 5개 데이터베이스 중에서 DIMD에서만 "도보라 참조"가 있는 것으로 나타났고, 다른 데이터베이스에서는 상호참조가 제공되지 않고 있다.

대소문자의 구별 실험결과 각 데이터베이스는 대소문자를 구별하지 않고 있으며 소문자로 검색을 실시한 경우에도 검색과정은 대문자로 인식하여 결과를 나타내고 있다.

4.3 완전성 평가

데이터베이스의 구성면에서 불완전성에 대한 사전안내나 범위의 구체적인 소개등의 여부와, 각 레코드의 필드구성, 그리고 접근 가능한 필드와 제한탐색등의 가능성, 비 논리적인 검색에 대

한 탐색의 결과등의 오류에 대해 실험 검색을 하였다. KAIST의 경우, 약식도움말의 지원만이 제시되고 있었으며 데이터베이스에 대한 사전 안내나 이용자 탐색지원을 위한 도구를 사실상 찾을 수 없었다. 필드 구성면에서는 간략필드로 5개 사항을, 상세필드로 9 내지 11개 필드사항을 이분시켜 선택하게 하고 있다. 접근이 가능한 필드는 8개로 KINITI나 KORDICT 보다는 상대적으로 약간 많은 접근필드를 제공하고 있다. 오류사항을 점검한 'X-ray' 용어의 검색에서는 〈표 7〉에서와 같이 자료형태별 검색결과와 총 검색결과와의 불일치가 보였다.

〈표 8〉에서 보듯이 주제범위의 특성에 따라 KORDIC의 각 데이터베이스는 다양한 필드구성을 갖추고 있으며 필드의 구성면에서 KISS, TECTREND는 비교적 통일성이 있으나 UNION과 SATURN는 자료의 성격에 따라 필드의 구성을 달리 하고 있다. 그러나 필드가 생략되

〈표 7〉 KAIST 데이터베이스 완전성 평가표

DB명	불완전성의 언급	필드구성	접근점 필드와 제한탐색	'무명씨' 검색등 오류사항
KAIST	약식도움말만 제공, DB주제, 구성에 대한 것 없음	간략 5필드 검색번호, 서명, 저자명, 출판사, 년도 상세 9-11필드 청구기호, 저자명, 회의명, 서명, 출판사항, 형태사항, 총서명, 일반주기, 기타주기, 미국목록번호, ISBN, 언어, 주제명, 학위논문정보	서명, 저자명, 출판사, 주제명, 보고서, 등록번호, ISBN, LCCN 8	'X-ray' 주제의 DT별 검색결과와 총수 불일치 단행본-213 전집-0 연간물-2 보고서-16 석사논문-35 박사논문-11 동주제에 대한 총합 (215)과의 차이 53건

〈표 8〉 KORDIC 데이터베이스 완전성 평가표

DB명	불완전성의 언급	필드구성	접근점 필드와 제한탐색	'무명씨' 검색등 오류사항
UNION	알림정보 DT에 대한 불분명한 설명	12 필드 동서: 제어번호, 저자명, 서명, 출판사항, 형태사항, 소장기관 양서: LCCN, ISBN, LC청구기호, DDC 청구기호, 일반주기, 주제명	저자명, 서명, 단체명, 회의명, 개인명, 기본색인필드, 6필드	DT필드 데이터누락
SATURN	없음	9필드 자료유형, 언어, 소장정보, 제목, 저자명, 형태사항, 초록, 주제명, 원문수록처, 입력일자, 제어번호	서명, 저자명, 주제명, 초록, 원문수록처 5	초록필드 거의생략
KISS	없음 directory service 제공	8필드 한글서명, 영문서명, 한글초록, 영문초록, 한글저자명, 한글공저자, 권호, 페이지범위	한글서명, 한글초록, 영문서명, 영문초록, 저자, 권호 (6)	영문서명이나 초록, 공저자명 데이터누락
TECTREND	없음	10필드 제어번호, 작성일자, 호수, 제목, 참고자료, 참고자료 수록일, 주제분류, 본문	서명, 본문, 참고자료, 주제분류, 출판년도, 권호수	논문자료만 디렉토리서비스로 최신성 유지

거나 표에 기술된 필드 이외에 주제코드, 판차사항, 학위논문주기, 서지주기, 연속간행물 등의 필드가 특수 지시나 알림사항 없이 추가되어 있는 경우가 많다.

각 데이터베이스별로 검색접근 필드를 비교해 보면, UNION과 SATURN은 필드 구성이 다양한 반면에 이용자가 접근할 수 있는 필드는 소수

로 제한되어 있다. 특히 검색시 가장 큰 어려움이었던 것은 출판년도에 따른 검색이 불가능하여 최신의 서지정보에 대한 연도별 검색이 지원되지 않았고, 언어별이나 문헌 형태에 따른 검색이 불가능했다. 검색접근 필드가 극히 제한되어 있는 것은 데이터베이스 구성에 있어서 검색위주의 구성이라기보다 저장위주의 구성에 치우쳐 있음을

〈표 9〉 KINITI 데이터베이스 완전성 평가표

DB명	불완전성의 언급	필드구성	접근점 필드와 제한탐색	'무명씨' 검색등 오류사항
BIST	없음, 단 서지가 없는 것은 생략으로 알림	10필드 자료형태, 기사구분, 사용언어, 제작년월, 제목, 저자명, 서지사항, 분류, 초록, 색인어	DT, LA, YR, 기사구분만 재검색가능	DT별 검색결과 불일치 (7건차이) J - 1029 T - 029 C - 60 E - 0
INFO	없음	10필드 자료형태, 기사구분, 사용언어, 제작년월, 제목, 저자, 서지사항, 분류, 초록, 색인어		초록 거의 생략 J - 102 C - 1 T - 0 E - 1 (총합과 0건차이)
DIGS	없음	9필드 년도, 문헌번호, 주제코드, 저자, 기사제목, 잡지명, 서지사항, 주제분류, 색인어	재검색은 연도만 가능	'무명씨' 로 저자명필드 접근으로 2건 검색결과 나옴
DIMD	없음	8필드 년도, 문헌번호, 학위코드, 저자, 대학, 논문제목, 발표일, 주제분야, 색인어, 초록	재검색은 학위코드로만 가능	초록 거의 생략 D - 40 M - 440 (총합과 차이 0건)
KSMA	없음	DT, LA, TI, AU, CS, SO, CC, DE, AB	재검색은 학위코드로만 가능	J - 339 C - 0 T - 0 E - 0

나타낸다고 할 수 있다.

KISS와 TECTREND는 레코드 상에 있는 모든 필드가 검색될 수 있으며 논문과 동향지의 리스트를 권 호별로 제공하는 디렉토리 서비스를 지원하는 장점을 갖고 있다. 각 데이터베이스는 기본적으로 자료형태, 기사구분, 사용언어, 제작년월, 제목, 저자, 서지 사항, 분류, 초록, 색인어의 형식을 가지고 있으며 이 가운데 데이터베이스의 특성에 따라 몇 가지 필드가 생략되었다. DIGS는 초록이 없었고, DIMD는 학위논문에 관한 구성요소 필드를 하나 더 포함하고 있었다. BIST와 INFO는 필드 구성요소가 일치했으며, KSMA는 BIST의 구성요소를 영어 알파벳 약자로 표현하고 있다.

〈표 9〉의 검색결과를 살펴보면 DIMD와 INFO는 자료형태별 검색결과 각 검색문서의 합이 전체와 일치하는 것으로 나타났으며, BIST와 KSMA에서는 결과가 일치하지 않았다. 각 필드별 검색결과와 건수의 합과 총 문헌수의 비교에서 결과의 차가 가장 큰 것은 KSMA로 자료형태 필드에서 J자료만 검색되었으며 다른 형태는 결과가 0건으로 검색되었다. 이것은 시스템 자체 내에서 J이외의 자료형태를 인식하지 못하거나 'OR' 연산자가 제 기능을 못하는 것으로 볼 수 있다. 특별히 무명씨 단어에 대한 검색은 저자명의 오류를 미리 예상하고, 오류가 될 수 있는 단어를 실험 검색한 것인데, 검색결과 DIGS에서만 두 가지 문서가 저자 명에 무명씨가 기입되어 검색되었다.

5 결 론

5.1 요약

과학기술정보 데이터베이스의 품질을 평가하기 위해 정확성, 일관성, 완전성의 3가지 평가기준에 따라 국내의 과학기술 관련 주요 기관인 연구기관, KORDIC, 산업기관, KINITI, 그리고 학술기관인 KAIST 등의 세 기관 데이터베이스들을 개별적으로 실험 검색하였고, 그에 대한 각각의 검색결과를 종합, 분석했다. 본 연구에서 제시한 연구문제에 따라 분석된 결과를 요약하면 다음과 같다.

연구문제 1) 과학기술정보 데이터베이스의 전자정보로서의 품질은 어떠한가?

정확성 평가에서 KORDIC의 경우, 전 실험대상 데이터베이스에서 외래어 표기상의 부정확성이 발견되었고 특히 UNION 과 SATURN에서의 철자오류나 KISS의 초록 필드 상에서의 영어 단어의 철자오류가 발견되었다. KINITI의 경우에도 실험대상 5개 전 데이터베이스에서 외래어 표기 상 오류가 많이 발견되었다. 대다수의 오류는 자료형태나 언어 등 특정필드의 데이터가 누락되어 있는 경우이며, 국문초록필드에서는 문법 상 맞지 않는 문장이 있거나 표기 상 철자오류가 있었고 DIGS에서는 연도 접근시 논리 값이 아닌 수치로 검색결과가 산출되는 오류를 나타냈다. KAIST의 과학도서관 데이터베이스의 경우는 외래어의 부정확성으로 각기 달리 표기된 용어에 따라 다른 검색결과를 주고 있으며 OR 연산 결과와도 일치하지 않으며 국문 초록필드 상에서 문법에 맞지 않고 뜻이 전달이 안 되는 문장이 나타났다. 실험 시 정확성 측정기준에 따라 상기한 데이터베이스들은 많은 오류를 보이고 있어 품질 면에서 적합하다고 보기 힘들다.

일관성 평가에서 연구대상인 10개의 데이터베이스는 유사한 결과를 보이고 있다. KORDIC의

경우에, 주제별 상호참조서비스는 전혀 제공되고 있지 않으며 약어처리나 대소문자 구별은 동일하게 취급하고 있는 반면 띄어쓰기는 각각 다른 용어로 인식하고 있어 사용자가 탐색 시 이에 대한 안내가 없이는 정확성이 높은 검색결과를 갖기 힘들다. 용어 'BASIC' 과 'basic' 의 경우와 같이 대소문자가 각기 다른 의미를 갖는 경우나, 어떤 용어의 띄어쓰기가 이용자의 실수로 달라지는 경우에 이를 통제할 수 있는 방법이 없다. 특히 SATURN의 경우에 약어인 DB가 레코드 상에서 실체는 여러 번 등장함에도 불구하고 약어 DB를 DATABASE로 인식하지 못하고 있어 검색결과는 0건으로 나타나고 있다. KINITI도 약어나 대소문자처리, 띄어쓰기에서 같은 결과를 보였으며 DIMD를 제외한 모든 데이터베이스에는 상호참조가 없었다. KAIST의 경우에 주제별 상호참조는 없으며 약어를 인식은 하나 동일한 용어로 처리하고 있지 않으며 대소문자 구별이 없고 띄어쓰기가 다른 경우에 각각을 다른 용어로 인식하고 있다. 일관성 측정기준에 따라 상기한 연구대상에 대한 실험결과는 오류를 보이고 있다.

완전성 평가에서 KORDIC의 4개 데이터베이스는 각 데이터베이스 자체에 대한 사전안내나 알림정보를 제공하지 않고 있다. KISS의 경우에만하여 디렉토리 서비스를 제공하고 있으며 이용자가 검색 전에 알아야 할 내용이나 필드구성, 혹은 데이터 입력 시 나타날 수 있는 누락정보나 공백필드에 대한 안내는 전혀 없다. KINITI의 5개 데이터베이스들의 경우도 자체 불완전성에 대한 사전언급은 없으며 2차 검색 시에 접근점으로 한정시켜 놓은 필드가 지극히 적어 재현율이 아닌 정확률을 높이기 위한 제한탐색이나 'NARROW-DOWN' 검색이 거의 불가능했다. 특히 저자 명 필드에 '무명씨'로 검색을 실시한

결과, DIGS에서는 2건의 검색결과가 나타났다. 또한 특정 주제어를 선정하고 이에 대해 자료형태별 검색을 실시한 결과와 모든 형태의 자료를 검색한 결과가 일치하지 않는 경우가 BIST와 KSMMA에서 나타났다. KAIST는 이용자 도움말이라는 약식의 안내만 제공하고 있으며 데이터베이스 자체의 구성이나 잠재적인 불완전성, 혹은 있을 수 있는 오류에 대한 처리 등에 대한 사전언급은 없다. 특히 과학도서관의 장서에 대한 주제 안내가 미약했으며 특히 DB내에 교수연구논문은 포함되고 있지 않다는 안내는 도움말이나 초기검색화면 어디에서도 찾아 볼 수 없었다. 특정 주제용어를 선정하고 자료형태별로 실시한 검색과 모든 자료를 망라하여 검색한 검색결과는 일치하지 않았으며 도움말로 제공되는 내용이 1997년으로 남아 최근갱신이 안되고 있음을 알 수 있었다. 종합적으로 볼 때 완전성 측정에서 상기한 연구대상들은 적지 않은 오류와 미비점을 보이고 있다.

연구문제 2) 데이터베이스 정보자체의 품질 측정에서 지적된 문제들은 무엇이 있는가?

엄청난 분량의 데이터가 다양한 조건과 형식에 따라 무 오류의 완전한 데이터베이스로 구축되는 것은 불가능할 수 있다. 의미 없이 저장되는 데이터가 살아있는 중요한 정보로 변화되는 정보시스템은 그 자체가 유기적으로 항상 변화되어야만 하는 구조이기 때문이다. 그러나 중요한 것은 그 때문에 데이터베이스는 오류수정이나 정보 자체의 품질에 대한 평가가 그 어떤 것보다도 절대 필수적이다.

본 연구의 실험검색을 통하여 분석한 결과에서 한가지 중요한 사실은 KORDIC이나 KINITI, 혹은 KAIST 그 어느 것도 특별히 월등하거나 특정

적으로 우수한 측면이 없이 거의 유사한 수준의 정보품질을 갖고 있다는 것이다. 특히 오류발생 측면이나 각 측정기준면에서 미비한 부분들이 전반적으로 비슷하다는 것이다. 다음은 연구자가 실험검색을 통하여 분석한 결과로부터 지적할 수 있는 문제점들이다.

가) 데이터베이스 구축 시에 데이터베이스 자체에 대한 공식적인 설명이나 안내가 누락되어 있다.

본 연구자는 다음과 같은 공식적인 안내내용을 최소한의 정보로 나열한다:

- 1) 데이터베이스가 다루고 있는 주제범위에 대한 상세한 분류와 이에 제외되고 있는 영역에 대한 명확한 안내, 그리고 관련 인접학문 영역에 대한 소개,
- 2) 제작진과 저작권 사항,
- 3) 데이터베이스 구축과 제작 년도,
- 4) 데이터베이스 전체 구조와 레코드 구성,
- 5) 데이터 갱신주기와 관련된 정보,
- 6) 레코드상의 누락필드나 공란정보의 해설,
- 7) 수록되고 있는 원천자료에 대한 출판 년도를 비롯한 상세한 정보,

나) 검색되는 색인어나 주제어의 용어통제가 완벽하게 이루어지지 않고 있다.

특히 외래어의 경우는 표기상의 오류가 많고 동일한 용어임에도 불구하고 약어처리나 대소문자구분, 띄어쓰기에 따라 각각 다른 검색결과를 주고 있다. 이것은 다 주제 인접학문이나 첨단과학과 같은 최선의 새로운 주제일 경우에 외래어나 신조어의 처리에 오류를 가중시키게 되어 검색접근측면에서 치명적인 검색오류를 야기시킬 수밖에 없다. 제 기능을 할 수 있는 시소러스가

절대적으로 요구된다.

다) 최종 이용자 검색을 지원하고 있는 색인어 브라우저이나 주제별 상호참조와 같은 서비스가 데이터베이스내의 데이터 검색을 지원하지 못하고 있다.

최소한 첨단 과학기술분야의 연구소나 산업지원정보센터, 혹은 학술기관이라면 과학기술분야의 특수한 정보 성격에 맞추어 보다 효과적인 데이터베이스 이용자 검색지원 서비스도 지원이 되어야 한다. 그러나 연구대상의 경우에는 검색 시 기본 연산자 적용에 대한 약식 도움말 이외엔 없었다. 네트워크에 의한 최종이용자 검색환경에서 기본적인 이용자검색지원이 제 기능을 하고 있다고 보기 힘들다.

라) 검색 시 발생할 수 있는 오류와 그에 대한 이용자의 오류처리안내나 설명이 검색과정에서 제공되지 않거나 실질적인 도움을 주지 못하고 있다.

본 연구자는 검색을 지원하는 이용자 서비스의 기본적인 사항을 아래와 같이 제시한다:

- 1) 검색메뉴와 자료에 대한 안내,
- 2) 검색 접근점으로 적용할 수 있는 탐색필드와 검색 시 오류정보의 제공,
- 3) 이용자서비스로 제공되고 있는 색인어 브라우저이나 각종 서비스에 대한 안내,
- 4) 특정분야에서 효과적인 검색단계나 검색전략에 대한 지침서나 지시문,
- 5) 잠재적으로 발생될 수 있는 오류처리에 대한 이용자 지시문,
- 6) 검색 시 도움말자료와 관련성이 높은 참고자료나 다른 데이터베이스에 대한 소개.

마) 과학기술 분야의 특정학문으로 갖고 있는 정보의 특성을 감안하여 이에 따른 보다 효과적인 검색기능의 엔진을 갖추는데 투자를 하거나 그에 따른 검색전략개발이나 전문적인 시스템 개선책 등이 전체 연구대상기관에서는 활성화되지 않고 있다.

오히려 기존의 데이터베이스에 대한 지속적인 개선과 향상을 위한 투자나 노력을 하기보다는 이와는 또 다른 시스템, 예를 들면 전자도서관이나 디지털 도서관등의 완전히 기존의 것과는 연계가 안되거나 기존의 도서관 환경과 어우러지기 힘든 별도의 새로운 시스템을 구상하거나 구축하는데 예산과 노력을 지나치게 기울인다면 이점을 문제로 볼 수도 있다. 기존에 이미 제작되어 사용되어 온 데이터베이스와 그를 검색하는 시스템에 대한 개선적인 측면의 투자는 그에 대한 검토와 평가 그리고 계속적인 개선이 투여되지 않는 한 거의 손실로 여겨질 수 밖에 없으며 그에 따른 엄청난 국가전체의 예산상, 정책상 낭비는 더 이상 받아들여질 수 없는 사실이다.

연구문제 3) 국내 과학기술연구 정보의 데이터베이스 품질개선을 위해 제시될 수 있는 최우선적인 정책이나 방안은 무엇인가?

첫째, 데이터베이스의 공공성이 강조되는 정보화 사회에서는 데이터베이스 생산자나 제작자들이 준거할 수 있는 품질 평가 기준 안이 공식적으로 마련되어야 한다. 무엇보다도 데이터베이스 생산자나 제작자에 의한 우선적이고 일반적인 정보 품질 관리방법은 데이터베이스 구축 시 공식적인 설명서를 작성하도록 정책적으로 요구하거나 국내 데이터베이스 제작 연합회나 국제기관과 긴밀한 협력관계를 유지하고 국제 학계 심포지움 등에서 실시하는 전문가수준의 교육을 실시하거

나 참여하게 하여 데이터베이스 제작 시 공식적인 사전안내에 대한 절대 필수적인 측면을 각성시켜야 할 것이다. 이것은 국내제작 데이터베이스의 품질을 국제적인 측면에서도 인정받을 수 있도록 하는데 필수 불가결한 작업이며 가장 우선적으로 해결해야 할 문제이다.

둘째, 기존에 생산된 데이터베이스의 품질 수준을 사회적, 제도적으로 명시하고 어떤 혜택을 부여하여 적극적으로 권하거나 제작시 필수사항으로 요구하는 제도가 반드시 필요하다. 정보품질에 대한 공식적인 기준과 그에 의해 평가된 데이터베이스를 품질 수준에 따라 차별화 함으로써 데이터베이스 품질에 대한 사용자 인식을 높이고 데이터베이스를 개발하는 제작자나 연구자들의 노력을 유도한다.

셋째, 데이터베이스의 품질을 심사하는 권위 있는 전문기관의 필요성을 인지하게 하고 국가단위의 공식적인 기구로서 각 기관이나 도서관 혹은 데이터베이스 관련 조직들의 상위기관으로 국내제작 데이터베이스에 관한 한 품질 평가와 관리를 총괄하는 기능을 갖게 한다. 제작된 데이터베이스의 주제분야별로 전문가들을 포함하며 데이터베이스 검색과 관리에 경험이 많은 전문가들을 구성하여 평가대상이 되는 데이터베이스를 일정기간 실제 검색을 통하여 평가하고 그에 대한 평가자문과 개선책을 공개하도록 한다.

넷째, 데이터베이스를 관리하거나 유지하는 시스템은 철저한 전거통제 기능을 수행하는 시스템을 갖추어야 한다. 생산된 데이터베이스의 정보 일관성을 유지시키는데 기본적인 방안중 하나가 시소러스의 개발이다. 이에 대한 지속적인 투자가 있어야 하며 데이터의 수집, 색인, 입력, 검증 등을 지원하는 것도 시소러스의 구축에 의하여 지원될 때 비로소 적절한 품질관리가 이루어질

수 있다. 특정 분야의 데이터베이스가 생산될 때 그 제작진에 의한 시소러스구축은 데이터베이스 구축과 병행될 때 큰 가치를 기대할 수 있으나 기존 국내 사정으로는 일반적으로 차후에 시소러스 구축이 진행되고 있다. 비록 그러한 경우라도 데이터의 품질 관리가 효율적으로 유지되기 위해서는 시소러스개발을 보다 앞당기는 정책이 반드시 이루어져야 한다.

다섯째, 데이터베이스가 일단 제작이 되고 이의 유지 관리가 시작되면 지속적으로 되어야 하는 작업이 오류 수정 소프트웨어나 자동관리시스템을 적용하여 지속적인 데이터 평가를 실시하는 것이다. 예를 들어 스캐닝, 자체 점검 데이터, 데이터 자동검증(automated data validation) 등은 데이터베이스의 오류수정이나 정확성 등을 검토하는 효과적인 방법이라 할 수 있다. 새로운 데이터의 구축만큼이나 필수적인 데이터 검증이 별도의 전문가에 의하여 제도적으로 이루어져야만 한다. 따라서 궁극적으로는 지능적인 데이터베이스 품질관리시스템이 어느 기관이나 필수적으로 유지될 필요가 크다. 이러한 시스템은 데이터 자체에 대한 전반적인 검토뿐만 아니라 데이터 내용에 대한 전문가 수준의 검토를 수행하고 자동적인 오류 점검 기능을 수행하는 것이다.

여섯째, 데이터베이스를 유지하고 관리하는 정보서비스 전문가들의 정보자체 품질에 대한 보다 신중한 관심과 평가기술이 훈련되어야 하며 제도적으로 이를 분석하고 평가할 수 있는 능력을 배양시켜야 할 것이다. 데이터베이스도 인공으로 제작된 것이며 여러 단계에 걸쳐 원천정보가 가공된 것이니 만큼 데이터 품질에 대한 분석자의 자세와 지속적이고 정확히 평가하는 노력을 정보 전문가들이 강조해야 한다.

5.2 제언

국내에서 과학기술로 제작되고 서비스되는 데이터베이스 10개를 대상으로 실험검색을 실시하는 과정에서 본 연구자와 본 연구자 지도하에 본 연구의 실험탐색자료를 이차자료로 석사논문에서 먼저 사용한 연구원들이 공통적으로 얻은 사실 몇 가지를 소개한다.

우선은 자료 검색 시 발생한 문제나 문의사항에 대해서 직접 해당 기관을 방문하지 않는 한 전화나 전자우편으로는 피드백을 받기가 거의 불가능했다는 점이다. 검색자의 도움을 얻고자 할 때 생각보다 시간이 많이 걸렸고 정확한 응답이 오지 않거나 여러 차례 문의를 해도 두 달이 지나도록 아무런 피드백이 없는 경우도 있어 해당기관에서 네트워크를 통한 이용자 검색지원을 하고 있다고 보기가 어려웠다. 이러한 상황은 네트워크환경에서 국내 대표적 과학기술관련 정보관리와 그 서비스가 이루어지고 있는가에 대해 의심하게 했으며, 이것이 인력의 부족이 원인인지 정보전문가로서의 전문성이 문제인지 아니면 그 두 가지의 복합적인 것이나 다른 것이 문제인지를 철저히 분석할 필요성을 절실히 느꼈다.

연구대상으로 선정된 기관들 중에는 전자도서관 프로젝트와 유사한 사업들을 진행하는 경우가 있었고 나름대로 정보서비스측면에 개선점을 강구하려는 면도 없지 않았다. 그러나 전체적으로 볼 때 앞서서도 지적한 대로 기존 시스템을 유지하고 제 기능을 최대한 발휘하기 위해 진단과, 철저한 평가, 혹은 신중한 개선책 모색을 위한 심도 있는 연구를 지속적으로 지원하기보다는 또 다른 사업으로 새로운 관심과 투자가 계속 이전되고 있는 문제를 지적할 수 있다.

마지막으로 과학기술 분야에서 최선의 주제나

새로운 연구를 수행하는 연구자나 전문가들이 이 용자로서 얼마나 각 해당기관에서 제공하고 있는 정보를 신뢰하고 기대하고 있는가를 철저히 조사할 필요가 있다. 국제 정보에 의존하는 풍토는 지구촌이 되고 있는 정보산업사회에서는 당연하다고 할 수 있다. 그러나 국내 과학기술 연구집단이 국내에서 제작된 데이터베이스의 정보 품질면에 불신이 크거나 국내 제작 데이터베이스로 부터는 가치 있는 정보를 얻기 힘들다고 여기고 있다면

문제는 심각하다. 국내에서만 수집이 가능하고 제작될 수 있는 한국의 고유정보는 외국을 상대로 할 때 엄청난 자산이므로 최소한 국내의 고유 정보에 대한 정보품질유지와 관리만이라도 철저히 이루어져야 한다. 데이터베이스의 가치를 유지하고 정보로 활용하기 위해서는 정보품질 측정과 그에 따른 평가는 없어서는 안될 정보전문인의 필수적인 작업이다.

참 고 문 헌

- 김기수. 1996. 자료의 질 향상을 위한 데이터베이스의 최적 감사 시점. 『데이터베이스저널』, 3.
- 김선형. 1997. 『과학기술DB의 품질평가에 관한 연구』. 석사학위논문 서울여자대학교 대학원 문헌정보학과
- 김엽영. 1992. 『온라인 데이터베이스 정보유통의 활성화 방안 연구』. 석사학위논문. 성균관대학교 행정대학원.
- 김정규. 1991. 『한글 문자 데이터베이스의 품질 평가 방안에 관한 연구』. 인천대 공업개발연구원.
- 김지훈. 1994. 서지데이터베이스의 품질관리. 『도서관학 논집』, vol. 21.
- 남영호, 김치용, 조만형. 1994. 『2000년대를 향한 국가 과학기술정보 유통시스템 구상』. 과학기술정책관리연구소.
- 박수경. 1996. 『정보시스템 컨설팅 서비스의 품질 요인에 관한 연구』. 석사학위논문. 서강대학교 경영대학원.
- 박 석. 1996. 멀티미디어 데이터베이스 시스템의 현황. 『데이터베이스월드』, 39.
- 송영희. 1995. 국내 데이터베이스 산업 현황과 활성화 방안에 관한 연구. 『동의논집』(인문·사회과학), 22.
- 유사라. 1997. 『하이퍼미디어 도서관 정보시스템』. 한국도서관협회.
- 유혜영. 1997. 『국내 제작 데이터베이스의 평가에 관한 연구』. 석사학위논문. 서울여자대학교 대학원 문헌정보학과.
- 이국희. 1996. 데이터베이스 품질에 관한 연구. 『데이터베이스 저널』, 3(1).
- 이원영. 1997. 『정보시스템 서비스에 대한 품질평가가 정보시스템 효과에 미치는 영향에 관한 연구』. 석사학위논문. 서울대학교 대학원 경영학과.
- 이용봉. 1996. 데이터베이스 품질에 관한 비평적 평가. 『국회도서관보』, 33(4).
- 이주현. 1995. 데이터베이스 표준화 중·장기 계획 수립을 위한 방향정립. 『데이터베이스월드』, 27.
- 한국데이터베이스진흥센터. 1994. 『데이터베이스 검색 표준화를 위한 기초연구』.

- . 1996. 『데이터베이스 표준화 연구보고서』.
- . 1996. 『데이터베이스구축 및 유통진흥사업에 관한 연구』.
- . 1997. 『97 데이터베이스 백서』.
- Amstrong, C. J. 1994. Database Quality: How users can help - And Be Helped! *In the Computers in Libraries International '94 Proceedings of the 8th Annual Conference*, Feb. 1994.
- Basch, Reva 1990. "Measuring the Quality of Data: Report and the fourth Annual SCOUG retreat." *Database Searcher*, 6(8).
- . 1995. *Electronic Information Delivery: Ensuring Quality and Value*. England: Gower Publishing Ltd.
- Bawden, D. 1990. *User-Oriented Evaluation of Information Systems and Services*. England: Gower Publishing Ltd.
- Dolan, D. R. 1992. "Quality Control at the System Level." *Online*, 16.
- Barrett, E. 1989. *The Society of Text: Hypertext, Hypermedia and the Social Construction of Information*. Cambridge: The MIT Press.
- Fox, C. et al. 1994. "The Notion of Data and its Quality Dimensions." *Information Processing & Management*, 30(1)
- Hary, V. & Oppenheim, C. 1993. "Evaluations of Electronic Database, Part I: Criteria for Testing CD-ROM Products." *Online & CD-ROM Review*, 17(4)
- . 1993. "Evaluations of Electronic Database, Part II: Criteria for Testing CD-ROM Products." *Online & CD-ROM Review*, 17(6)
- Jasco, P. 1992. *CD-ROM Software, Dataware and Hardware: Evaluation, Selection, and Installation*. Englewood, Colorado : Libraries Unlimited, Inc..
- Liepins, G. E. & V. R. R. Uppuluri. 1990. *Data Quality Control: Theory and Pragmatics*. New York, NY: Marcel Dekker.
- O' Neill, E. T. & D. Vinzine-Goetz. 1988. "Quality Control in Online Database." *Annual Review of Information Science and Technology (ARIST)*, 23.
- Redman, T. C. 1992. *Data Quality: Management and Technology*. New York, NY: Bantam Books.
- Tenopir, C. 1987. "Online Database: Quality Control." *Library Journal*, 112.(3)