

과학기술분야 데이터베이스의 품질향상을 위한 품질평가 연구*

A Study of the Quality Evaluation for Improving the Database Quality in Scientific and Technical Fields

이 응 봉(Eung-Bong Lee)**
조 현 양(Hyun-Yang Cho)***
류 범 중(Beom-Jong You)****
최 재 황(Jae-Hwang Choi)*****

목 차

1. 서 언	지침, 시소러스 유지관리 및 오류수정
2. 연구의 목적 및 방법	4. 4 메뉴검색, 키워드검색 및 복수 데이 터베이스 검색에 대한 기능 분석
3. 과학기술분야 데이터베이스 품질평가 기준	4. 5 실제의 검색과정 및 검색결과를 통한 문제점 분석
3. 1 선행연구 분석	4. 6 검색식에 소요되는 평균 검색시간
3. 2 선행연구 분석을 통한 품질평가 기준 설정	4. 7 화면 디자인에 대한 평가
3. 3 데이터베이스 품질측정 평가항목	4. 8 검색결과와 디스플레이 및 출력
4. 데이터베이스 품질평가	4. 9 데이터베이스 이용안내
4. 1 잘못된 데이터 값이 발견된 레코드	4. 10 Help Menu 및 Q & A
4. 2 레코드 필드구조	5. 결론 및 제언
4. 3 레코드수 차이, 갱신주기, 레코드 작성	

초 록

본 연구는 객관적으로 검증 가능한 데이터베이스 품질평가 기준 및 방법을 개발하고 과학기술분야 데이터베이스를 대상으
로 한 실증적인 실험연구를 목적으로 하고 있다. 평가기준은 데이터베이스 수록데이터 및 데이터베이스 제공서비스의 두가지
측면을 고려하여 전체 18개의 평가항목을 설정하였으며, 데이터베이스의 품질평가는 연구개발정보센터(현, 한국과학기술정보
연구원)에서 구축하여 인터넷의 웹 상에서 이용자 서비스를 제공하고 있는 우리나라 과학기술분야의 대표적인 7개의 데이
터베이스를 대상으로 하였고, 평가결과를 종합·분석하여 문제점을 도출하고 향후의 개선방안을 모색하였다.

ABSTRACTS

This study aims to evaluate the quality of databases in scientific and technical fields. The main discussion is about the
quality issues of KORDIC's seven databases, the most typical web database of scientific and technical fields in Korea. Eighteen
evaluation criteria, considered two aspects of data itself and service of database, are used to measure their quality. The
purpose of this study is to diagnose the quality of databases mentioned above and to analyses the measured result of quality
objectively and to find out the problems of them and also to provide possible suggestions for their improvements.

키워드: 데이터 품질, 데이터베이스 품질, 데이터베이스 서비스 품질, 품질평가 기준

- * 본 논문은 2000년도 연구개발정보센터(현, 한국과학기술정보연구원)의 지원에 의하여 이루어짐.
** 충남대학교 사회과학대학 문헌정보학과 조교수(eblee@cnu.ac.kr),
*** KISTI 정보유통사업실장, **** KISTI 전자정보개발실장, ***** KISTI 정보유통사업실 선임연구원
접수일자 2001년 5월 26일

1. 서 언

인터넷을 통한 정보 공유가 가능해지면서 이용자들은 정보 공유의 필요성과 중요성을 더욱 실감하고 있다. 또한 기업에서 제공하고 있는 정보 서비스에 대한 이용자의 만족도는 기업과 조직의 이미지만이 아닌 실제 기업의 성장에 중요한 요소가 되고 있다.

인터넷이 보급되기 전 이용자들은 대부분 정보를 찾기 위해 서점, 도서관 등의 장소에 직접 방문해 정보를 찾는 것이 대부분이었다. 그러나 최근 몇 년 동안 인터넷의 발전과 개인용 컴퓨터(Personal Computer; 이하 PC)의 급속한 보급으로 인하여 정보를 찾는 방법이 변화되었다. 과거에는 서점, 도서관을 직접 방문해서 직접 목록을 확인하여 정보를 찾았다면, 인터넷이 보급된 후에는 웹 상에서 서지 목록이나 초록을 검색한 후, 서점이나 도서관에서 정보를 찾게 되었다. 또한 인터넷 사용이 일반화되고 인터넷에서 제공하는 정보가 폭넓고 깊어지면서 굳이 서점이나 도서관을 찾지 않아도 인터넷에서 정보검색을 하고 웹 상에서 전문 정보(Fulltext Database)까지 이용하는 것이 가능하게 되었다. 따라서 이전보다 더욱 더 신속하고 효율적으로 정보를 찾을 수 있게 된 것이다.

최근 몇 년 동안 급속히 증가하는 인터넷 정보수요를 충족시키기 위해 정부는 물론 일반 데이터베이스 제공자는 질보다는 양적인 데이터베이스 구축에 힘을 기울였다.

사실 데이터베이스의 품질 문제는 어제, 오늘의 일이 아니라 컴퓨터가 탄생한 이래 줄곧 이어져 온 고질적인 문제라 할 수 있다. 그러

나 최근에 이르러 이 문제가 더욱 심각해지기 시작한 것은 데이터베이스의 사용이 보편화, 대중화 추세에 접어들면서부터이다. 인터넷 등 통신망의 확산과 상용 또는 공공 데이터베이스의 개발 및 보급의 확산으로 인해 데이터베이스 사용자층이 종전의 전문(expert) 사용자로부터 PC 앞의 초보(naive) 사용자 또는 일반대중으로 이동하게 되었다. 따라서 일반 사용자들은 불량 데이터의 위험에 거의 속수무책이 되고 말았다. 전문 사용자는 불량 데이터를 예방하거나 발견하는 나름대로의 경험을 가지고 있으며, 스크린에 나타난 데이터가 어떻게 입력이 되었는지 알고 있으므로 그 정확도가 어느 정도 될 것인지를 한 번쯤은 생각한다. 그러나 일반 사용자는 스크린에 나타난 데이터를(그것이 명백한 철자오류가 아니라면) 100% 믿어 버리는 경향이 있다. 쉽게 믿어버리는 만큼 그로 인한 피해의 빈도(frequency)와 심도(severity)는 커질 수 밖에 없는 것이다.

일반 사용자가 겪는 이러한 피해는 단순히 개인적인 문제가 아니라 광의로는 정보사회의 성공 여부, 협의로는 데이터베이스 산업의 사활에 영향을 미친다고 할 수 있다. 우리나라의 경우 많은 수의 데이터베이스가 급속히 구축되고 있지만 데이터베이스의 현행화(update) 등 유지보수에 대한 대책이 마련되지 않는다면 대부분 데이터베이스의 품질은 몇 년 내로 형편없는 수준으로 곤두박질할 수 밖에 없을 것이다. 호기심으로라도 이러한 데이터베이스에 접근해 온 사용자는 저품질(low quality) 데이터에 실망하게 되고, 한번 실망한 사용자는 두 번 다시 비싼 사용료를 내고 사용하지 않

을 것이며, 점점 사용자층이 메말라 가는 데이터베이스 산업의 수익성은 더욱 나빠질 것이다. 또한 열악한 수익성으로 인해 민간투자자와 고급인력의 유입은 더욱 어려워지며, 지속적인 투자가 없는 한 데이터베이스의 현행화 및 발전적 갱신(enhancement)은 커녕 기본적인 유지보수 활동마저 어려워 데이터베이스의 스크레기화는 더욱 가속화되고 악순환은 반복될 것이다.

따라서 데이터베이스 품질문제는 국내 데이터베이스 산업이 성장(또는 생존)하기 위해 넘어야 하는 높은 문턱으로 남아 있는 셈이다. 이제까지의 데이터베이스 산업 육성지원 전략이 “규모의 확대와 양적 증가, 그리고 사용자층의 확산”에 있었다면 이제부터의 전략은 “데이터베이스의 품질향상을 통한 질적 발전 그리고 사용자층의 만족 제고”에 초점을 맞추어야 한다. 그러기 위해서는 데이터베이스 품질문제에 관한 철저한 분석과 체계적인 연구가 선행되어야 한다. 또한 인터넷이 보급되기 전 데이터베이스의 개념과 인터넷의 보급으로 인해 웹 상에서 데이터베이스 품질기준 또한 달라져야 한다. 기존 데이터베이스와는 다르게 현재 웹 상에서의 데이터베이스는 다른 특징적인 것들을 요구하고 있다.

특히, 국내외에서 생산되는 과학기술분야 연구물의 홍수 속에서, 우리 나라의 과학기술분야의 데이터베이스를 평가하여, 개선 방안을 모색하는 것은 시급한 일이 아닐 수 없다. 이러한 데이터베이스의 지속적인 품질향상은

연구자들이 필요로 하는 데이터의 탐색과 수집에 소요되는 노력과 시간을 단축하고, 실험과 분석에 보다 많은 시간을 투자하게 함으로써, 양과 질적으로 우수한 연구물을 생산과 직결되는 것이다. 또한 효율적인 정보환경을 조성함으로써, 관련분야 연구인력의 연구생산성을 증대하고, 연구자들이 사이의 동일 연구과제에 대한 중복연구를 방지하여 국가 연구생산성을 향상시킬 수 있다. 이러한 질적인 데이터베이스의 발전은 결국 이용증진과 신뢰할 수 있는 과학기술 데이터베이스를 구축·서비스하는 것을 가능하게 할 것이다.

2. 연구의 목적 및 방법

21세기 지식기반사회로의 이행과 더불어 과학기술정보인프라구축의 중요성이 증대하면서 과학기술정보의 생산과 유통의 활성화를 위한 각종 정책 등이 강화되고 있다. KORDIC¹⁾(연구개발정보센터, 이하 KORDIC이라 한다.)는 국가차원의 과학기술전용망인 연구망의 고도화 및 과학기술정보유통을 위한 중추적인 역할을 수행하며, 이를 위해 과학기술정보 종합 데이터베이스를 구축·서비스하고 있다.

한편, 우리나라 대부분의 과학기술분야 데이터베이스는 質(질)보다는 量(양) 위주의 가공과 빈약한 전문인력에 의한 구축으로 질적 발전을 이루지 못하고 있는 실정이다. 또한 과학기술분야의 데이터베이스를 전문적으로 평

1) 연구개발정보센터(KORDIC)는 산업기술정보원(KINITI)과 2001년 1월에 통합하여 현재에는 한국과학기술정보연구원(KISTI)이란 이름을 사용하고 있다. 본 연구에서는 기존의 KORDIC 데이터베이스를 평가대상으로 선정하였기 때문에, 현재 KORDIC이라는 기관의 명칭은 존재하지는 않지만, 이해의 편의를 위하여 KORDIC이란 명칭을 사용하기로 한다.

가할 수 있는 공식적인 평가기준과 방법의 부재로 문제의 심각성이 더하고 있다. 이로 인해 수많은 예산과 노력의 낭비, 연구자들의 정보이용률 및 연구생산성의 저하를 초래할 것이다. 따라서 과학기술분야 데이터베이스 품질평가 기준 및 방법 개발을 통한 데이터베이스의 품질향상 및 과학기술정보의 질적 향상의 도모는 시급한 일이 아닐 수 없다.

본 고에서는 연구의 목적을 달성하기 위해 데이터베이스 품질평가에 관한 이론적 고찰 및 선행연구의 비교분석을 병행하여 실시하였다. 이론적 연구에 있어서는 기존의 서적과 간행물, 관련문헌, 학위논문, 기타 연구 발표된 자료들을 토대로 데이터베이스 품질에 관한 개념을 체계적으로 정리하였고, 선행연구에서는 해외 데이터베이스 품질평가와 국내 데이터베이스 품질평가에 관한 연구를 각각 비교함으로써 각각의 차이점과 선행연구에서의 한계점을 도출하였다. 또한 실증분석에서는 신뢰성 분석과 요인분석을 사용하여 새로운 웹 기반 환경에서 가장 설명력 있는 데이터베이스 품질평가 기준들에 관하여 분석을 수행하였다. 이러한 결과를 토대로 국내 데이터베이스의 품질 실정에 적합한 새로운 데이터베이스 품질평가 기준을 제시하였다. 그리고 제시한 품질평가 기준을 적용하여 KORDIC의 데이터베이스를 평가하고, 평가된 결과를 토대로 KORDIC 데이터베이스의 품질 문제의 해

결을 위한 발전방향을 제시하였다.

3. 과학기술분야 데이터베이스 품질평가 기준

3.1 선행연구 분석

데이터베이스를 평가하는데 있어서 가장 중요한 문제는 데이터베이스를 올바르게 평가할 수 있는 기준의 설정이라 할 수 있다. 데이터베이스를 어떻게 평가할 것인지를 결정하고, 평가결과에 대한 정확성 및 적합성을 분석하기 위해서는 몇 가지 평가이론들이 전제되어야 할 것이다. 본 연구에서는 데이터베이스 품질평가에 관한 여러 평가기준 및 이론들에 대한 종합·분석·평가를 통하여 본 연구에서 적용하고자 하는 데이터베이스의 평가기준을 설정하고자 한다. 그리고 데이터베이스 품질평가에 관한 평가기준 및 이론들에 대한 종합·분석·평가대상 선행연구는 국내 3편 및 국외 14편 총 17편의 관련 논문 및 보고서를 대상으로 하였다.²⁾

〈표 1〉의 분석결과에 나타난 바와 같이 여러 연구에서 데이터베이스 품질을 파악할 수 있는 기준들을 제시하고 있다. 데이터의 정확성 같은 일부 기준은 대부분의 연구에서 등장하고 있다. 하지만 어떤 기준은 여러 연구에서

2) 본 연구에서 분석대상으로 선정한 데이터베이스 품질평가 기준과 관련된 연구는 다음과 같다. 1. 한국데이터베이스진흥센터의 데이터베이스 품질평가에 관한 연구, 2. 이재환의 분산체제로 구축된 통합 데이터베이스의 품질검증 및 개선을 위한 연구, 3. 박준식의 색인의 평가기준에 관한 분석적 연구, 4. SCoug, 5. Peter Jasco, 6. Brodie, 7. Doll & Torkzadeh, 8. AT & T Bel, 9. Herther, 10. Large, 11. Granick, 12. Information Automation Limited and CIQM, 13. 콜로라도 데이터베이스 네트워킹 콘소시움의 일반 데이터베이스 평가법, 14. SCOUT Report, 15. Greg R. Notess, 16. 뉴욕대학 법학스쿨의 범 데이터베이스, 17. 온라인 데이터베이스 사용자 그룹에서 채택된 데이터베이스

서로 다른 의미로 사용되기도 하고, 상이한 용어로 표현되는 기준이 실상 동일한 의미를 지니는 경우도 있다. 특히 데이터 자체의 품질 부분보다는 CD-ROM의 기능과 특성을 나타내는 서비스 품질 부분 기준들이 더욱 통일이 되어 있지 않음을 볼 수 있다.

그리고 데이터베이스 품질 평가에 관한 연구들이 대부분 데이터 자체 품질에만 치중하고 있는 반면 최근에 중요시되고 있는 서비스 측면이 소홀히 되고 있는 것을 볼 수 있다. 인터넷 환경이 급속도로 발전하면서 웹을 기반으로 한 데이터베이스의 서비스가 급증하고

있는 추세를 감안한다면 기존의 연구에서는 이와 같은 최근의 정보환경을 충분히 반영하고 있지 않는 것으로 나타났다.

〈표 2〉에 나타난 바와 같이 국내·외 선행 연구 17편의 품질평가 기준은 총 16개의 품질기준으로 요약할 수 있다. 정확성, 최신성, 완전성 및 수록범위 등은 여러 연구에서 중요한 품질기준으로 제시하고 있는 반면에, 유일성이나 하드웨어 등은 특정 연구에서만 품질기준으로 제시하고 있는 것을 볼 수 있다.

그리고 대부분의 연구가 데이터베이스 시스템 구축 및 서비스, 이용자 측면, 데이터베이

〈표 1〉 국내·외 데이터베이스 품질평가 기준 비교 분석

품질기준	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	비 고
수록범위		○	●	●	●					●	●	●		●	●		●	● 9 01
정확성	●	○	●	●	○	○	●	●	●	●	●	●		●	●	●	●	● 13 03
일관성	●	●	○	●		○	○				●	●		●		●	●	● 8 02
완전성	●	●	●		●	●	●	●				●			●	●	●	● 11 00
최신성	●	●	●	●	●		●	●	●	●	●	●		●		●	●	● 14 00
전문성										●						●		● 2 00
통합성				●				●			●	●		●				● 5 01
권위			●		●				●	●				●		○		● 5 00
검색성	●				○											●		● 2 01
접근성					●		●				●					●		● 4 00
사용이 용성	●			●			●		●		●	●		●		○	●	● 8 01
사용자 지원성	●			○	●		●		○		●	●	○	●			●	● 7 02
비용			●	●	●		○				●	●	●	●			●	● 8 01
네트워크				●	●				●				●					● 2 00
하드웨어				●	●							●	●					● 3 00
기 타		자체 제작 비율		이용자 호응도		신속성				유일성		이용자 호응도						

- 각 연구에서 품질평가 기준으로 평가되었던 항목
- 각 연구에서 품질평가 기준은 아니지만 고려되었던 항목

스 자체 품질 등에 대한 통합된 평가기준이 아닌 단편적인 평가기준만을 제시하고 있으며 특히, 국내의 품질평가 연구는 외국에 비해 전반적인 연구가 미흡한 상황이다. 최근 들어 품질평가에 대한 논의가 일고 있지만, 큰 성과를 이루지 못하고 있는 실정이다.

또한 데이터베이스 유형을 고려하지 않고, 일반 데이터베이스의 품질평가 기준을 적용하고 있다. 실제로 최근 한국데이터베이스진흥센터의 “데이터베이스 이용실태 및 정보 수요조사 결과” 보고서에 의하면 이용자들의 40% 이상이 웹을 통한 정보 이용시 정보의 정확성보다는 용이성 및 접근성을 가장 큰 정보 이용의 중요성으로 지적하고 있는 것으로 나타났다.

3. 2 선행연구 분석을 통한 품질평가 기준 설정

데이터베이스의 품질은 데이터베이스의 바람직한 정도 또는 우수성이라고 정의할 수 있으며, 다음과 같이 데이터 품질과 서비스 품질로 구성이 된다고 할 수 있다.

데이터베이스 데이터 품질이란 데이터베이스가 수록하고 있는 데이터 자체의 바람직한 정도와 데이터 양 그리고 전문성을 의미하며 다음과 같은 다섯가지 차원에 의해 분석될 수 있다.

- 데이터 구조(Structure)
- 데이터 값(Value)
- 데이터 표현(Representation)
- 데이터 양(Capacity)
- 데이터 전문성(Expert)

그리고 데이터베이스 서비스 품질이란 데이터가 사용자에게 서비스되는 과정에서의 매체, 수단, 방법, 그리고 공급처 지원 등의 바람직한 상태를 의미하며 다음과 같은 다섯가지 차원에 의해 분석될 수 있다.

- 서비스의 정확성(Accuracy)
- 서비스의 편리성(Convenience)
- 서비스의 친절성(Kindness)
- 서비스의 신속성(Promptness)
- 서비스 비용(Cost)

따라서, 본 연구에서는 데이터베이스의 품질평가에 적용될 품질기준을 데이터베이스 데이터 품질과 데이터베이스 서비스 품질의 두

〈표 2〉 국내·외 데이터베이스 품질평가 기준 사례연구 분석

데이터베이스 품질평가 기준			
수록범위	9	검색성	2
정확성	13	접근성	4
일관성	8	사용용이성	8
완전성	11	사용자지원성	7
최신성	14	비용	8
전문성	2	네트워크	2
통합성	5	하드웨어	3
권위	5	기타	5
합계			106

가지 측면을 포함하여 <표 3>과 같은 10가지의 데이터베이스의 품질기준을 제시한다.

3. 3 데이터베이스 품질측정 평가항목

본 연구에서는 데이터베이스의 품질을 데이터베이스 데이터 품질과 데이터베이스 서비스 품질에 관한 측면 양자를 동시에 고려하고, 앞에서 제시한 선행연구 분석을 통한 데이터베이스 품질평가 기준을 참고로 하여 <표 4>와 같은 데이터베이스 품질측정 평가항목을 마련하였다.

4. 데이터베이스 품질평가

과학기술분야 데이터베이스의 품질향상을 위한 품질 평가기준으로 설정한 18개의 평가항목에 의거하여, KORDIC에서 구축하여 운영 중인 7개의 데이터베이스 즉, 과학기술전문정보 데이터베이스, 과학기술도서정보 데이터베이스, 석·박사학위논문 데이터베이스, 국내학회논문정보 데이터베이스, 연구보고서 데이터베이스, 과학기자재 데이터베이스 및 과학기술동향 데이터베이스에 대한 품질을 평가하였다.

평가작업은 각각의 데이터베이스 내에 포함되어 있는 레코드 전반에 대한 분석, 각각 데

<표 3> 데이터베이스 품질평가 기준

구분	품질기준 (Criteria)	핵심관점
데이터 품질	정확성 (Accuracy)	데이터베이스 데이터가 실제 값과 동일한가?
	완전성 (Completeness)	표현하고자 하는 실세계의 중요한 객체들과 속성들이 담겨있는가?
	최신성 (Currentness)	가장 최근의 데이터로 갱신되었는가?
	수록범위 (Record Scope)	데이터베이스에 수록되는 저널의 종수와 수록기간은 언제까지인가?
	전문성 (Expert)	해당분야를 얼마나 포함하고 있으며, 학술적 가치가 있는 자료를 얼마나 가지고 있는가?
서비스 품질	검색성/접근성 (Searching)	검색이 얼마나 신속하게 그리고 정교하게 이루어지는가?
	사용 용이성 (Ease of Use)	인터페이스를 통한 데이터베이스접근과 산출정보 활용이 얼마나 쉽고 편리한가?
	사용자 지원성 (Customer Support)	Documentation, Training, Help 등 사용자 지원이 적합한가?
	비용 (Cost)	비용은 얼마나 소요되며, 데이터베이스 이용 비용은 적정한가?
	네트워크 및 하드웨어 (Network & Hardware)	데이터베이스 이용 속도는 어떠한가?

〈표 4〉 데이터베이스 품질측정 평가 항목

평가기준	평가항목	Poor	Fair	Good
정확성	1. 레코드 내에 잘못된 데이터 값이 어느 정도 발견됩니까?			
완전성	2. 레코드의 필드 구조는 완전합니까? 그리고 동일시스템에서 제공하는 서로 다른 DB의 필드 및 구조에 일관성이 있습니까?			
완전성	3. 공표된 레코드 수와 실제 레코드 수가 차이가 있습니까?			
최신성	4. 얼마나 자주 데이터를 갱신합니까?			
정확성/완전성	5. 레코드 작성과 관련하여 신뢰할만한 기준이 있습니까?			
검색성	6. 시소러스의 유지관리가 원활히 이루어지고 있습니까?			
사용자 지원성	7. 오류 데이터 수정은 체계적으로 이루어지고 있습니까?			
검색성	8. 다양한 탐색기법을 제공하고 있습니까?			
사용 용이성	9. 메뉴 검색의 경우 메뉴 구조의 체계성이나 메뉴에서 제공되는 검색경로의 다양성이 어느 정도 우수하고 정교합니까?			
사용 용이성	10. 키워드 검색은 사용이 용이하고 다양합니까?			
검색성	11. 복수의 DB를 대상으로 검색이 가능합니까?			
검색성	12. 1건의 검색식을 검색하는 데 걸리는 평균 시간은 어느 정도입니까?			
사용 용이성	13. 화면디자인이 얼마나 이용자 지향적입니까?			
사용 용이성	14. 검색결과가 내용과 주제에 따라 다양하게 형태로 제공됩니까?(Display)			
검색성	15. 원하는 형식으로 검색결과와 출력이 가능합니까?			
사용자 지원성	16. DB 이용안내			
사용자 지원성	17. Help 메뉴			
사용자 지원성	18. Q & A			
	19. 기타			

데이터베이스의 운영자와의 면담 그리고 최근의 정보이용 환경이 인터넷의 Web을 통하여 이루어진다면 점을 감안하여 KORDIC의 기관 홈페이지는 물론 각각의 데이터베이스에 접속하여 검색실험을 실시하였고, 이용자적인 측면에서 정보이용과 관련된 항목에 관하여 심층적인 분석 및 평가를 실시하였다.

4. 1 잘못된 데이터 값이 발견된 레코드

4. 1. 1 철자의 오류

철자의 오류에 관한 평가는 일반적으로 대부분의 데이터베이스에서 잘못 입력되고 있는

10개의 영문단어 즉, bussiness, conditons, fransisco, government, grammer, infomation, misssion, recieve, retrieval, seperate를 탐색어로 설정하여 검색을 실시하였다.

〈표 5〉에 나타난 바와 같이 철자의 오류는 국내학회논문정보 데이터베이스에서 353건의 철자오류가 나타났고, 석·박사학위논문 데이터베이스에서 168건, 과학기술전문정보 데이터베이스에서 151건이 발생하는 등 평가 대상 7개의 모든 데이터베이스에서 발생하였다.

또한, 이와 같은 철자오류 뿐만 아니라, □ 표시가 나타나는 경우(예, Lef□vre, Herv□ 등) 및 글자가 깨져서 해독이 불가능한 경우

(예, 䄀ꄀꄀ이 ꄀꄀꄀ ꄀꄀꄀ ꄀꄀ ꄀꄀ 등)도 발견되었다. 그리고 평가 대상 KORDIC 데이터베이스들은 과거 여러번의 상호 통합과정을 거쳤기 때문에 시스템간의 데이터 변환시에 이와 같은 철자의 오류가 발생했을 개연성도 간과할 수 없는 사항이라 판단된다. 따라서 이와 같은 철자의 오류를 방지하기 위해서는 정기적으로 수록정보에 대한 검증작업을 실시하여야 함은 물론 철자 오류의 발견 시 이를 즉시 운영자에게 알릴 수 있는 메커니즘이 마련되어 바로 수정할 수 있는 방안을 구축하여야 할 것이다.

4. 1. 2 데이터의 불완전 기술

평가 대상 7개의 데이터베이스 공히 외국저자의 경우, 저자명이 완전하지 않은 형태로 나타나는 경우가 많이 있었으며, 불완전한 서명이 간혹 나타났으며 특히 과학기술전문정보 데이터베이스와 과학기술도서정보 데이터베이스의 경우, 해당자료의 소장기관 표시가 완전

하지 않은 경우가 빈번하게 발생하고 있는 것으로 나타났다(예, YSWL, KRIB, JNUL 등).

4. 1. 3 데이터 필드와 데이터 내용의 불일치
 데이터 필드와 데이터의 내용이 일치하지 않은 경우로는 빈번하지는 않지만 판차사항 필드에 발행년도가 표기되어 있다거나 ISBN 필드에 가격이 동시에 표기되어 있는 레코드가 발견되었다.

4. 2 레코드 필드 구조

레코드 필드 구조의 평가는 BI(Basic Index), TI(Title), KW(Keyword), AB(Abstract) 및 AU(author) 필드에 관하여 분석하였다.

〈표 6〉에 나타난 바와 같이 평가 대상 7개의 데이터베이스 중 레코드 필드 구조에 있어서 과학기술전문정보 데이터베이스, 국내학회 논문정보 데이터베이스, 석·박사학위논문정보 데이터베이스, 연구보고서 데이터베이스는

〈표 5〉 철자의 오류에 관한 검색결과

DB명 탐색어	과학기술 전문정보	과학기술 도서정보	석·박사 학위논문	국내학회 논문정보	연구 보고서	과학 기자재	해외과학 기술동향	계
bussiness	3	6	-	5	1	1	2	18
conditons	-	1	15	75	7	-	-	98
fransisco	7	-	-	-	-	-	1	8
goverment	45	28	8	34	9	-	-	124
grammer	88	19	11	4	3	-	-	125
infomation	5	-	8	12	2	-	3	30
misssion	2	-	-	-	-	-	-	2
recieve	-	1	3	13	1	1	-	19
retrival	-	2	2	23	5	-	-	32
seperate	1	6	121	187	50	8	1	374
계	151	63	168	353	78	10	7	830

4. 3 레코드수 차이, 갱신주기, 레코드 작성지침, 시소러스 유지관리 및 오류수정

본 항목에 대한 분석 및 평가는 해당 데이터베이스의 스키마를 비롯한 관련 자료, 각각의 데이터베이스 운영자와의 면담을 통하여 파악하였다. <표 7>은 본 항목에 대한 분석결과를 나타낸 것이다.

4. 3. 1 레코드 건수 차이

입력용 DBMS와 실제로 서비스되는 DBMS가 다르기 때문에 시간에 따른 차이는 평가대상 7개의 모든 데이터베이스에서 모두 존재한다고 판단된다. 그리고 과학기술전문정보 데이

터베이스, 과학기술도서정보 데이터베이스, 석·박사학위논문 데이터베이스 및 과학기자재 데이터베이스는 실제 레코드 건수와 서비스되고 있는 레코드 건수에 차이가 발생하지 않았다. 왜냐 하면, Edit를 이용하여 데이터베이스를 Text 형태로 작성한 후 KRISTAL II 데이터베이스에 업로드하는 형태이기 때문에 차이가 나타나지 않는 것이다. 그리고, 국내학회논문정보, 연구보고서 데이터베이스 또한 레코드 건수에 있어 차이가 나타나지 않았다.

4. 3. 2 갱신주기

갱신주기도 레코드 건수와 마찬가지로 각각의 데이터베이스가 사용하는 DBMS에 따라서

<표 7> 레코드수 차이, 갱신주기, 레코드 작성지침, 시소러스 유지관리 및 오류수정 분석

DB명 평가항목	과학기술 전문정보	과학기술 도서정보	석박사 학위논문	국내학회 논문정보	연구 보고서	과학 기자재	해외과학 기술동향
레코드 수 차이	No	No	No	Yes	Yes	No	Yes
갱신 주기	분기별	수시	년 1회	수시	수시	수시	매주
레코드 작성지침	Yes	MARC 사용	MARC 사용	Yes	Yes	Yes	Yes
시소러스 유지관리	No	No	No	No	No	No	No
오류 수정	No *	확인불가	수시	수시	수시	수시	수시
사용 DBMS	KRISTAL II	KRISTAL II	KRISTAL II	KRISTAL II	KRISTAL II	KRISTAL II	KRISTAL II
비고	입력 DBMS와 서비스 DBMS가 다름	-	-	-	-	입력 DBMS와 서비스 DBMS가 다름	입력 DBMS와 서비스 DBMS가 다름

* 현재의 시스템은 한번 수정 삭제 시 데이터베이스 자체를 재시동 해야함

차이가 나타나는 것으로 밝혀졌다.

4. 3. 3. 레코드 작성지침

평가 대상 7개의 데이터베이스 모두가 레코드 작성지침을 가지고 있는 것으로 나타났다.

4. 3. 4 시소러스 유지관리

평가 대상 7개의 데이터베이스 모두가 시소러스 자체를 구축하고 있지 않기 때문에 당연히 시소러스의 유지관리는 이루어지지 않고 있는 것으로 밝혀졌다.

4. 3. 5 오류수정

현재 과학기술전문정보 데이터베이스의 경우, 오류수정이 이루어지지 않고 있다. 왜냐 하면, 현재의 시스템 특성상 한번 수정·삭제 시 데이터베이스 자체를 재시동해야 하기 때문이다(KRISTAL II 데이터베이스의 특성상 이에 대한 모듈개발이 아직까지는 이루어지지 않고 있다). 그리고 나머지 평가 대상 6개의 데이터베이스들은 오류가 발견될 때마다 수시로 수정이 이루어지고 있는 것으로 나타났다.

4. 3. 6 사용 DBMS

평가 대상 7개의 데이터베이스 모두가

KRISTAL II를 사용하고 있다. 그리고 이 중 과학기술전문정보 데이터베이스, 과학기자재 정보 데이터베이스 및 해외과학기술동향 데이터베이스의 경우에는 입력 DBMS로는 ORACLE을 사용하고 있다.

4. 4 메뉴검색, 키워드검색 및 복수 데이터베이스 검색에 대한 기능 분석

<표 8>에 나타난 바와 같이, 평가 대상 7개의 모든 데이터베이스에서 키워드검색은 상세검색과 고급검색으로 나누어 각 필드를 지정하여 검색을 용이하게 지원하고 있고, 검색엔진이 각 데이터베이스에서 색인어를 추출하여 제시하고 제시된 색인어를 클릭하여 데이터를 찾을 수도 있으며, 복수 데이터베이스 검색을 지원하고 있는 것으로 나타났다.

그러나 메뉴검색의 경우, 과학기술전문정보 데이터베이스와 과학기술도서정보 데이터베이스 2개를 제외한 나머지 5개의 데이터베이스는 메뉴검색을 지원하고 있는 것으로 나타났다. 메뉴검색을 지원하는 5개의 데이터베이스는 각각의 데이터베이스 특성상 차이를 보이고 있다. 석·박사 학위논문 데이터베이스의 메뉴검색은 '디렉토리 열람'이라는 페이지

<표 8> 메뉴검색, 키워드검색 및 복수 데이터베이스 검색에 대한 기능 분석

DB명 탐색기법	과학기술 전문정보	과학기술 도서정보	석·박사 학위논문	국내학회 논문정보	연구 보고서	과학 기자재	해외과학 기술동향
메뉴검색	X	X	O	O	O	O	O
키워드 검색	상세검색	O	O	O	O	O	O
	고급검색	O	O	O	O	O	O
색인어검색	O	O	O	O	O	O	O
복수DB검색	O	O	O	O	O	O	O

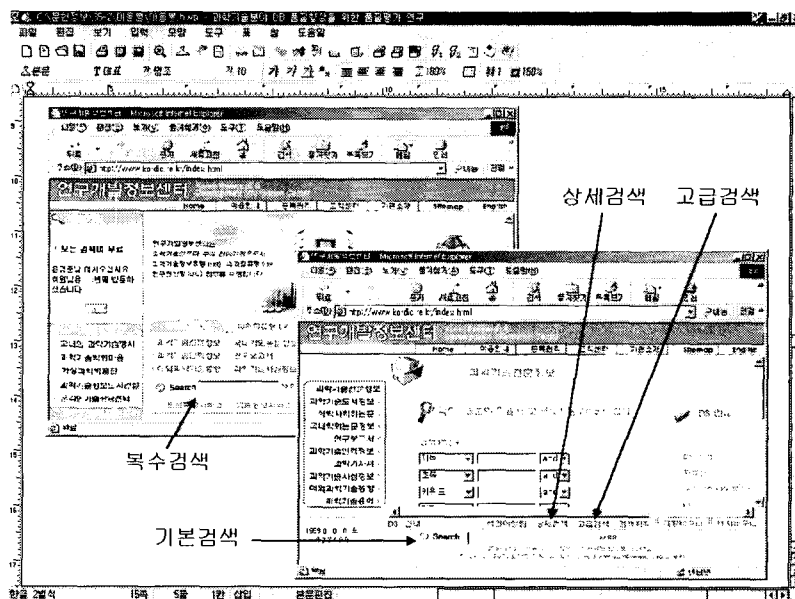
지를 두어 학교별로 학과를 구분하여 연도별로 분류를 하여 검색을 할 수 있고, 국내학회 논문정보 데이터베이스의 메뉴검색은 '디렉토리 열람'이라는 페이지를 두어 150개의 학회에 대한 논문을 학회지와 논문대회의 논문으로 구분하여 연도와 권·호수 별로 분류를 하여 검색을 할 수 있으며, 연구보고서 데이터베이스의 메뉴검색은 '디렉토리 열람'이라는 페이지를 두어 과학기술부, 정보통신부, 농림부 등 국가에서 지원한 연구과제의 연구결과보고서 원문을 부처별 연도순으로 분류하여 검색을 할 수 있고, 과학기자재 데이터베이스의 메뉴검색은 '디렉토리 열람'이라는 페이지를 두어 대학 또는 연구소 및 국·공립기관이 보유하고 있는 고가의 과학기자재를 지역, 기관명, 도입년도의 3단계로 분류를 하여 검색을 할 수 있으며, 해외과학 기술동향 데이터베이스의 메뉴검색은 '디렉토리 열람'이라는 페이지

를 두어 정보제공자들이 제공한 해외과학기술 동향정보를 20개의 대주제와 그에 따른 소주제, 년도의 3단계로 분류를 하여 검색이 가능하도록 하고 있다.

4. 5 실제의 검색과정 및 검색결과를 통한 문제점 분석

탐색기법을 활용한 실제의 검색과정 및 검색결과를 통한 문제점을 분석하기 위하여 <그림 1>에서와 같이 KORDIC의 기관 홈페이지 및 개별 데이터베이스에서 제공하는 검색기능의 종류를 복수검색, 기본검색, 상세검색, 고급검색의 네가지로 설정하여 각각에 대하여 다양한 탐색식을 사용하여 검색을 실시하였다.

실제로 'cell' 과 'structure' 라는 탐색어를 설정하고 구탐색(탐색어 사이에 스페이스를 한칸 두는 경우), AND(&), OR(!), NOT(!)



<그림 1> KORDIC 홈페이지 상에서 제공하는 검색 창의 종류

연산자를 사용하여 검색을 수행하였을 때의 각각에 대한 검색 결과를 분석하였다. 평가 대상 7개 데이터베이스에 대한 검색결과 분석

은 <표 9>와 같다. 평가 대상 7개의 데이터베이스에 대한 검색실험은 2000년 12월 1일부터 12월 7일까지 수행하였다.

<표 9> 실제의 탐색 결과 (단위: 검색된 레코드 건수)

탐색의 종류	DB명	과학기술 전문정보	국내학회 논문정보	석박사 학위논문	연구 보고서	과학 기자재	해외과학기술동향	과학기술 도서정보
cell structure	복수	6	34 s	32 s	4 s	-	1 s	18 s
	기본	6	34 s	32 s	238 s	-	17 s	18 s
	고급	-	-	-	-	-	-	-
	상세	-	-	-	-	-	-	-
cell & structure	복수	1,392	422 s	429 s	238 s	2 s	-	101 s
	기본	1,392	422 s	429 s	238 s	2 s	err	101 s
	고급	0	0	0	0	②	-	0
	상세	50,691 s	0	269 s	질의 err	②	0 ①	1 s
cell structure	복수	482,694	14,487 s	6,383 s	3,222 s	244 s	-	10,789 s
	기본	482,694	14,487 s	6,383 s	3,222 s	244 s	err	10,789 s
	고급	399,149	4,191	1,305	1,377	②	-	11
	상세	777,462 s	4,191	4,981 s	질의 err	②	962 s	69 s
cell ! structure	복수	45,447	4,037 s	1,336 s	1,127 s	196 s	-	2,409 s
	기본	45,447	4,037 s	1,336 s	1,127 s	196 s	err	2,409 s
	고급	399,148	4,189	1,305	1,377	②	-	11
	상세	③	③	③	③	②	-	③
cell & structure ! plant	복수	1,386	422 s	429 s	212 s	2 s	-	101 s
	기본	1,386	422 s	429 s	212 s	2 s	err	101 s
	고급	0	0	0	0	②	-	0
	상세	③	③	③	③	②	-	③

* <표 9>는 탐색어를{(cell structure), (cell& structure), (cell | structure), (cell ! structure), (cell & structure ! plant)}으로 설정하여 불 연산자를 사용하였을 경우의 실제 검색 결과치이다.
 * 고급 및 상세 검색에서는 '초록(AB)' 필드로 검색을 제한한다.
 * s: stemming을 나타낸다.
 * ①: 해외과학기술동향 데이터베이스는 AB가 초록이 아닌 본문으로 나타난다.
 * ②: 과학기자재 데이터베이스는 AB 항목이 존재하지 않는다.
 * ③: 상세검색에서는 NOT 연산자를 사용할 수 없다.
 * 해외과학기술동향 데이터베이스는 기본검색에서 연산자 지정이 안되며 고급검색이 지원되지 않는다.
 * 복수 데이터베이스 검색의 경우 에러가 빈번하게 발생하는데, 이를 해결할 수 있는 안내가 어디에도 존재하지 않는다.

4. 6 검색식에 소요되는 평균 검색시간

- System: Microsoft Windows 98 Second Edition

검색식에 따른 평균 검색시간의 측정은 다음과 같은 세가지의 방식을 사용하여 측정하였다. 첫째, 기본검색 창을 사용하여 단일어(가공) 및 &연산자 사용(정보 & 기술)의 두 가지 방식의 검색식을 작성하여 검색을 실시하였다. 둘째, 고급검색 창을 사용하여 검색식(TI: 네트워크 & TI:시스템)을 작성하여 검색을 실시하였다. 그러나 해외과학기술동향 데이터베이스는 고급검색을 지원하지 않기 때문에 제외하였다. 셋째, 검색은 평가 대상 7개의 데이터베이스에 대하여 5번씩 실시하였고, 각각에 대한 소요시간 및 평균값을 구하였다. 여기에서 검색식에 소요되는 검색시간이라 함은 검색식을 작성한 후 search key를 누른 후 검색결과 값이 나오기까지 걸리는 시간을 의미한다.

검색에 사용된 PC의 환경은 다음과 같으며, 네트워크는 연구전산망을 사용하였다.

- Processors: Intel Pentium III 501MHz
- Model: Information P6K(Katmai)Pentium III(B) 450-600 2-2.05V
- Installed Memory: 128MB RAM

<표 10>에 나타난 바와 같이 '가공'이라는 단일어를 사용한 검색식을 검색하는데 소요된 시간은 평가 대상 7개의 모든 데이터베이스에서 평균 1초 내외의 상당히 빠른 결과를 나타냈다.

<표 11>에 나타난 바와 같이 & 연산자를 사용한 '정보 & 기술'이라는 검색식을 검색하는 소요된 시간 역시 평가 대상 7개의 데이터베이스 공히 1초 내외의 시간이 소요되었다. 그러나, 해외과학기술동향 데이터베이스의 경우에는 &연산자가 실행이 안되고 띄어쓰기를 &연산자로 인식하기 때문에 검색식을 '정보 기술'로 수정하여 검색한 결과이다.

<표 12>에 나타난 바와 같이 고급검색 창을 사용하여 'TI: 네트워크 & TI: 시스템'이라는 검색식을 작성하여 검색을 실시한 결과 기본 검색 창을 사용한 검색결과와 마찬가지로 1초 내외의 매우 빠른 결과를 나타냈다. 이 경우, 해외과학기술동향 데이터베이스는 고급검색을 지원하지 않기 때문에 제외하였다.

<표 10> 검색식에 따른 검색 소요시간: 단일어 검색

(단위: 초)

DB명	검색시간					평균	검색건수	비고
	1회	2회	3회	4회	5회			
과학기술전문정보	1.12	0.97	1.11	1.47	1.13	1.16	2774건	탐색어: '가공'
과학기술도서정보	1.01	0.81	0.63	0.73	0.62	0.76	983건	
석·박사학위논문	0.71	0.87	0.83	0.99	0.84	0.85	638건	
국내학회논문정보	1.13	1.21	0.93	0.77	0.80	0.97	1812건	
연구보고서	0.83	0.74	0.83	0.86	0.86	0.82	1137건	
과학기자재	0.83	0.97	1.00	0.77	0.67	0.85	128건	
해외과학기술동향	1.12	0.69	0.76	0.67	0.95	0.84	2159건	

〈표 11〉 검색식에 따른 검색 소요시간: & 조합검색

(단위: 초)

DB명	검색시간					평균	검색건수	비 고
	1회	2회	3회	4회	5회			
과학기술전문정보	1.94	1.79	2.21	1.81	1.93	1.94	867건	검색식: '정보 & 기술'
과학기술도서정보	1.10	1.13	0.83	1.12	1.29	1.09	322건	
석·박사학위논문	0.99	0.89	1.03	0.97	0.93	0.96	43건	
국내학회논문정보	1.07	1.19	1.47	1.09	1.19	1.18	442건	
연구보고서	1.23	1.09	0.84	1.01	0.97	1.03	715건	
과학기술자재	1.76	0.93	0.62	1.07	0.77	1.03	33건	
해외과학기술동향	1.2	1.04	1.24	1.11	1.12	1.15	1389건	

〈표 12〉 검색식에 따른 검색 소요시간: 고급 검색

(단위: 초)

DB명	검색시간					평균	검색건수	비 고
	1회	2회	3회	4회	5회			
과학기술전문정보	2.02	1.97	0.83	0.93	1.57	1.46	45건	검색식: 'TI:네트워크 & TI:시스템'
과학기술도서정보	1.04	1.25	1.23	0.89	1.07	1.10	6건	
석·박사학위논문	1.29	0.83	0.89	0.83	0.72	0.91	5건	
국내학회논문정보	1.79	1.15	0.88	0.91	0.69	1.08	7건	
연구보고서	1.25	0.84	0.82	1.05	0.98	0.99	2건	
과학기술자재	0.76	0.70	0.53	0.63	0.68	0.66	2건	

4. 7 화면 디자인에 대한 평가

화면 디자인에 관한 평가는 평가 대상 7개의 전체 데이터베이스에 대하여 Navigation, Layout, Color, Typography, Icon 및 Browser의 호환성 등의 여섯 가지의 카테고리를 설정하여 평가를 수행하였으며 평가결과는 〈표 13〉과 같다.

4. 8 검색결과 디스플레이 및 출력

검색결과 디스플레이에 있어서 평가 대상 7개의 모든 데이터베이스 공히 검색결과는 리스트로 보여지는 간략보기와 전체 레코드를 볼 수 있는 상세보기를 지원하고 있는 것으로

나타났다. 그리고 결과 리스트에서 보여지는 디스플레이 항목은 변경이 불가능하며, 결과 리스트에서 보여지는 항목은 오름차순, 내림차순, 검색순 정렬을 지원하고 있으며, 키워드 검색의 상세보기에서는 입력한 키워드가 하이라이팅(붉은색)으로 표현된다.

검색결과 출력은 사이트 자체에 별도의 출력기능은 없으며 검색결과는 저장이나 원문 복사 신청의 공간을 제공하여 개인 PC로 다운 받을 수 있도록 하고 있다.

4. 9 데이터베이스 이용안내

데이터베이스 이용안내는 KORDIC의 기관 홈페이지에서 일괄적으로 "검색도움말"이라는

〈표 13〉 화면 디자인에 대한 평가결과

평가 기준		평가내용	비고
Navigation	Navigation rule 적용여부	Yes	Rule은 적용되었지만 전체적인 레이아웃에 맞지 않음
	위치에 대한 적합성	No	데이터베이스별 이동 메뉴는 좌측에 위치하여 이동과 view가 용이하나 검색관련 메뉴와 기타 설명부분 메뉴가 하단부에 위치하여 이용시 적합하지 않음
	사이트 전체의 일관성	No	일관성있게 검색 page를 첫 화면에 띄우지 못함(제각각임). 과학기술전문정보와 국내학회논문정보를 제외하고는 첫 화면에 검색 창이 없음
	사용자 위치 파악 여부	Yes	페이지 상단에 관련 데이터베이스명이 명시됨
	텍스트/그래픽 /로딩 속도의 적합성	Yes	양호함
	사용의 편의성	No	Sub Menu(검색관련 메뉴)가 하단부에 위치해 있어 시선집중이 되지 않음. 각기 다른 단계의 Navigation 메뉴
Layout	Grid 사용여부	Yes	Table로 처리
	일관성 유지	No	화면구성이 동일하게 표현될 수 있음에도 불구하고 각기 다른 형태로 표현
	정보에 따른 유연성	Yes	Page 크기가 고정되어 있지 않음
	시선처리	No	화면이 너무 아래로 치우침
Color	Scheme에 의한 배열	Yes	각 데이터베이스별로 색을 지정하여 일관된 스타일 추구
	Main 및 응용 컬러 적용 여부 및 효과	Yes	각 데이터베이스별로 색을 지정하여 일관된 스타일 추구
Typography	가독성	No	이어지는 문장의 경우 줄 간격이 너무 짧음. Font 크기 10, 영어 Font 에 관한 지정내용이 없음(한글만 있음)
	요소들의 시각적 대비 효과	Yes	테이블로 시각적 대비
	배열, 그룹핑	Yes	Data출력량이 고정되어 있음(10건)
	가이드라인	Yes	정해진 Form에 맞추어 Data출력
Icon	메타포(Metaphor)	No	표현한 바와 다른 내용(자료실, FAQ 등). 상세검색과 고급검색과의 차이점이 없음
	스타일의 일관성	Yes	양호함
	사이트 디자인과의 매칭	Yes	메뉴에 있어서 회색과 검은색의 Concept를 유지
브라우저의 호환성		Yes	일반적으로 익스플로러에 많은 초점이 맞추어짐

메뉴로 제공되고 있다. 이용안내에 대한 분석 및 평가결과는 다음과 같다.

첫째, 웹 페이지가 1024*768 모드에 적합하도록 구성되어 있기 때문에, 대부분의 일반적

인 웹 페이지의 화면구성에서 적용하는 800*600 화면 모드에서는 전체(좌측 관리자 연결 및 레코드건수, 최종갱신일 부분)화면이 나타나지 않는 문제점을 지니고 있다.

둘째, 내용을 본 후, 상위 메뉴로의 이동성이 떨어지는 현상이 나타난다. 상세 세부 메뉴에서 상위 메뉴로 이동할 경우, 웹 브라우저의 Back Key나 좌측 메뉴 부분의 검색도움말 Icon을 선택해야 하는 번거로움이 나타난다.

셋째, 검색도움말 메인 페이지에 있는 데이터베이스 소개와 이미지 뷰어 설치 메뉴의 경우, 다른 메뉴와는 달리 하위 메뉴가 없기 때문에 마우스를 메뉴 위에 위치시키기 전에는 링크가 되어있는지의 여부를 파악할 수 없다.

넷째, 각 데이터베이스별로 이용안내가 데이터베이스 검색 메인 페이지에 있는 것도 있고, 제공되지 않는 데이터베이스의 경우도 있는데 "검색도움말"이나 "이용안내"라는 하나의 메뉴로 주 메뉴에서 통일하여 제공하거나 또는 각 데이터베이스별로 BI나 상세검색, 고급검색, 디렉토리 검색, 색인내용 등이 상이하므로 해당 데이터베이스에 맞는 이용안내를 만들어 각 데이터베이스 검색 메인 페이지에서 제공하는 것이 효율적일 것이라고 판단된다.

다섯째, 데이터베이스 검색안내 메뉴의 경우, 페이지 길이가 너무 길기 때문에 스크롤하는데 상당한 불편함을 초래한다. 또한 페이지 상단으로 가는 Icon이 각 페이지의 내용이 달라지는 부분에 위치하면 스크롤을 줄일 수 있음에도 불구하고 데이터베이스 검색안내 페이지의 제일 하단부에만 위치해 있다.

여섯째, 데이터베이스 소개, 데이터베이스 안내 등과 같은 의미의 내용임에도 불구하고

다른 페이지로 링크가 되어 있으며, 또한 전혀 별다른 특별한 의미가 없이 사용되고 있다.

마지막으로, 통합검색의 경우에 있어서도 일괄적으로 모든 데이터베이스를 검색할 것이 아니라, KINITI(현, KISTI) 데이터베이스에서 제공하는 통합검색의 경우에서처럼 데이터베이스 선택 option을 제공하여 검색의 편의성의 도모하는 것이 바람직할 것으로 생각된다..

4. 10 Help Menu 및 Q & A

Help Menu와 Q & A에 대한 평가항목은 다음과 같이 설정하였다. 우선, Help Menu에 대한 평가항목은 초보이용자에서 전문이용자까지 다양한 수준의 이용자 요구를 수용할 수 있도록 Help 메뉴가 구성되어있는지의 여부, 온라인 상의 이용자 질문에 얼마나 신속하게 대답하는지의 여부 그리고, Help 메뉴 이외에 고객 서비스를 위한 별도의 전화나 전자메일이 준비되어 있는지의 여부 등의 세가지를 설정하였고, Q & A에 대한 평가 항목은 이용자의 질문 시 적극적이며 친절하게 대답해 주는지의 여부 및 Error Message에 대한 해결방법의 제공 여부 등의 두가지를 설정하였다. 평가 대상 7개 데이터베이스의 Help Menu 및 Q & A에 대한 분석 및 평가결과는 다음과 같다.

첫째, 담당자의 전화번호와 E-mail Address를 제공되고 있지만, 인터넷의 Web을 통하여 정보를 제공하는 경우에는 BBS(게시판)형태로 Q & A를 제공하는 것이 일반적이며 효율적일 것이라 생각된다.

둘째, Help Menu 항목은 검색화면 메인 페이지의 "검색도움말"이라는 메뉴로 구성되어

있다.

셋째, FAQ는 과학기술전문정보 데이터베이스(회원가입, 이용요금, 원문서비스)와 석·박사학위논문 데이터베이스(원문보기, 오류원인) 뿐이며, 나머지 평가 대상 5개의 데이터베이스는 FAQ를 마련하고 있지 않다.

넷째, Q & A는 해외과학기술동향 데이터베이스(자료실 메뉴 - 자료실이라는 메뉴와 별다른 관련이 없음)를 제외한 6개의 데이터베이스는 자체적으로 Q & A를 구성하고 있지 않다. 즉, 6개의 데이터베이스는 KORDIC 기관 홈페이지의 상단 부분에 있는 고객센터라는 메뉴를 통해 Q & A의 기능을 제공하고는 있지만, 그 기능은 상당히 미흡한 실정이다.

5. 결론 및 제언

우리나라 과학기술분야의 정보유통에 크게 이바지하고 있는 KORDIC(현, KISTI)에서 구축하여 운영하고 있는 데이터베이스 중 본 연구에서는 7개의 데이터베이스 즉, 과학기술전문정보 데이터베이스, 과학기술도서정보 데이터베이스, 석·박사학위논문 데이터베이스, 국내학회논문정보 데이터베이스, 연구보고서 데이터베이스, 과학기자재 데이터베이스 및 과학기술동향 데이터베이스 대하여 크게 데이터베이스 수록 데이터 및 데이터베이스 제공 서비스의 두 가지 측면에서 전체 18개의 평가 항목을 설정하여 평가를 실시하였다. 평가 대상 7개 데이터베이스의 품질평가에 관한 주요 분석결과의 내용을 요약하면 다음과 같다.

1. 데이터베이스에 수록된 데이터의 품질적

인 측면에서 보면 평가 대상 데이터베이스 공히 만족할 만한 수준을 보이지 않고 있다. 모든 데이터베이스에서 철자의 오류(Typographical Error)가 발견되었고, 데이터가 불완전하게 기술되어 있거나, 데이터 필드와 데이터 내용이 일치하지 않는 경우도 발견되었다.

2. 레코드 필드의 구조는 비교적 완전하게 구성되어 있으며, 전체적으로 보아 일관성을 유지하는 것으로 나타났다.

3. 데이터베이스에 수록된 레코드 건수와 실제로 서비스되고 있는 레코드 건수는 차이를 보이지 않았다 그러나 평가 대상 7개의 데이터베이스는 입력용 DBMS와 서비스되는 데이터베이스 DBMS가 다르기 때문에 시간에 따른 차이는 현실적으로 발생할 수 밖에 없다.

4. 갱신주기는 각각의 데이터베이스가 사용하는 데이터베이스 DBMS에 따라서 차이가 나타났는데, 석·박사학위논문 데이터베이스의 경우에는 년 1회로 최신성에 문제점을 지니고 있으며, 수시로 갱신을 한다고 밝힌 데이터베이스조차도 과학기자재 데이터베이스와 해외과학기술동향 데이터베이스 제외한 나머지 4개의 데이터베이스는 최근 6개월 이내에는 한번도 갱신을 하지 않은 것으로 나타났다.

5. 레코드 작성지침은 7개 데이터베이스 공히 가지고 있었다. 그러나 시소러스를 구축하고 있는 데이터베이스는 하나도 없는 것으로 나타났다.

6. 과학기술전문정보 데이터베이스를 제외한 평가 대상 6개의 데이터베이스에서는 데이터의 오류에 대한 수정작업은 수시로 이루어지고 있는 것으로 나타났다. 그러나 과학기술전문정보 데이터베이스의 경우에는 오류수정

이 이루어지지 않고 있는 것으로 나타났다. 왜냐하면, 현재의 KRISTAL II 시스템의 특성상 한번 수정이나 삭제를 할 경우에는 데이터베이스 자체를 재시동 해야 하기 때문인 것으로 판단된다.

7. 과학기술전문정보 데이터베이스 및 과학기술도서정보 데이터베이스를 제외한 나머지 5개 데이터베이스 공히 메뉴검색을 지원하고 있으며, 복수 데이터베이스 검색 및 키워드 검색은 평가 대상 7개의 데이터베이스 공히 지원하고 있는 것으로 나타났다.

8. 실제로 탐색식을 작성하여 복수 데이터베이스 검색, 기본검색, 상세검색, 고급검색을 수행한 탐색결과에서는 많은 문제점을 가지고 있는 것으로 나타났다. 우선, 대부분의 데이터베이스에서 Stemming을 지원하고 있지만 사전에도 수록되지 않은 철자의 오류인 경우에도 Stemming이 되는 등 이에 대한 개선이 시급한 실정이다. 그리고 조합검색의 경우에는 NOT 명령어가 일률적으로 적용되지 않는 현상이 나타났다. 이와 함께 탐색어와는 전혀 상관이 없는 엉뚱한 탐색결과가 발생하는 등 검색시스템은 전반적으로 개선해야 할 문제점이 다수 있는 것으로 나타났다.

9. 검색식에 소요되는 평균검색 시간은 단일어 검색, & 조합검색 및 고급검색에서 평가 대상 7개의 데이터베이스 모두에서 평균 1초 내외의 시간이 소요되는 것으로 나타나 매우 양호한 상태를 나타냈다.

10. 화면의 디자인은 비교적 양호한 편이지만, 이용자 지향적인 측면에서는 개선해야 할 점이 많은 것으로 판단된다. 우선, Navigation의 경우에는 위치에 대한 적합성, 사이트 전체

의 일관성 및 사용의 편의성이, Layout은 일관성과 시선처리의 문제가, Typography와 관련된 사항으로는 가독성이, Icon과 관련된 사항으로는 메타포(metaphor)에 문제점을 가지고 있는 것으로 판단된다.

11. 검색결과를 디스플레이하는 경우에는 간략보기와 전체 레코드를 볼 수 있는 상세보기를 지원한다. 그러나 디스플레이 항목을 이용자가 원하는 방식으로 변경할 수는 없다. 또한 검색결과에 대한 사이트 자체의 출력기능은 없지만, 검색결과는 저장공간을 제공하여 이용자의 개인 PC로 다운로드 받을 수 있도록 하고 있다.

12. 데이터베이스에 대한 이용안내는 KOR-DIC의 기관 홈 페이지에서 일괄적으로 "검색도움말"이라는 메뉴로 제공되고 있다. 그러나 각각의 데이터베이스별로 이용안내가 메인 페이지에서 제공되는 것도 있고 제공되지 않는 데이터베이스도 존재한다. 또한, 데이터베이스 소개나 데이터베이스 안내 등은 동일한 의미임에도 불구하고 다른 페이지로 링크가 되어 있으며, 특별한 의미가 없이 사용되는 것으로 나타났다.

13. 평가 대상 7개의 데이터베이스 공히 Help Menu 기능을 위하여 담당자의 전화번호와 E-Mail Address가 제공되고 있지만, FAQ(또는 Q & A) 기능은 과학기술전문정보 데이터베이스(회원가입, 이용요금, 원문서비스)와 석·박사학위논문 데이터베이스만이 제공하고 있으며, 나머지 평가 대상 5개의 데이터베이스는 마련하고 있지 않은 것으로 나타났다.

이와 같이 우리 나라 과학기술분야의 종합

적인 정보유통기관으로서 중요한 역할 담당하고 있는 KORDIC에서 구축하여 운영하고 있는 7개 데이터베이스에 대한 품질평가의 결과는 만족스럽지 못한 일면을 보여주고 있다. 즉, KORDIC에서 구축하여 제공하고 있는 데이터베이스는 그 다양성과 축적된 레코드의 양적인 면에 비하여 품질적인 측면에서는 아직까지 보완되고 수정되어야 할 요소가 많은 것으로 나타났다

지금까지 우리 나라의 데이터베이스 관련 정책은 “규모의 확대와 양적인 증가 및 사용자층의 확산”에 치중하여 온 것이 사실이다. 그러나 앞으로의 데이터베이스 관련 정책은 “데이터베이스의 품질 향상을 통한 질적인 발전과 사용자층의 만족 제고”에 초점을 맞추어야 한다는 것을 염두에 두어야 할 것이다.

따라서 향후, 데이터베이스의 품질 향상을 통한 이용자의 만족도를 제고시키기 위해서는 아래와 같은 사항들이 고려될 수 있을 것이다.

첫째, 데이터베이스에 수록될 또는 수록된 데이터의 품질을 향상시키기 위해서는 현재의 운영방식인 데이터베이스의 분산구축방식을 중앙기관에 의한 철저한 감독 및 관리가 가능한 통합구축방식으로 전환해야 할 것이고 이와 함께 수록 데이터에 대한 검증과 오류 데이터의 수정을 정기적으로 시행할 수 있는 장치를 시급히 마련해야 할 것이다.

둘째, 오늘날과 같이 정보환경이 급변하는 상황에서는 정보의 적시성이 매우 큰 중요성을 가진다고 할 수 있다. KORDIC 데이터베이스의 경우, 데이터베이스의 갱신주기가 너무 길기 때문에 해외과학기술동향 데이터베이스

를 제외한 나머지 데이터베이스에서는 이용자들이 최신의 과학기술정보를 이용하는데 심각한 불편을 초래하고 있다. 따라서, 최소한 1달에 한번 정도는 데이터베이스의 내용을 갱신할 수 있는 체제를 반드시 구축해야 할 것이다.

셋째, KRISTAL II 검색엔진에 대한 개선이 시급히 이루어져야 할 것이다. 실제로 탐색식을 작성하여 복수 데이터베이스 검색, 기본 검색, 상세검색, 고급검색을 실시하여 보면 불리언 연산자 특히, NOT 명령어가 일률적으로 적용되지 않는다는지, Stemming의 무원칙성, 탐색어와는 전혀 다른 엉뚱한 용어로 탐색이 이루어지는 현상이 빈번하게 발생하고 있다. 오늘날, 데이터베이스 이용자들은 인터넷의 확산으로 국내·외의 성능이 우수한 검색엔진에 익숙해져 있다. 특히 KORDIC 데이터베이스를 이용하는 이용자들의 수준을 감안할 때, 검색엔진의 성능저하는 결국 이용율의 급감을 초래할 것은 명약관화한 일일 것이다. 따라서 검색엔진의 성능개선에 많은 노력을 기울여야 할 것이다.

넷째, 오늘날 데이터베이스의 이용은 인터넷상의 웹을 통하여 대부분 이루어지기 때문에 홈페이지의 성패는 내용은 물론 사용의 편의성에 달려 있다 해도 과언이 아닐 것이다. KORDIC의 홈페이지 중 데이터베이스 이용과 관련된 부분에서는 Navigation, Layout, Typography, Icon 처리 등 뿐만 아니라, 데이터베이스 이용안내와 FAQ(Q & A) 처리 등이 매우 미흡한 실정이다. 따라서, 이용자 지향적인 측면에서의 홈페이지에 대한 개선이 이루어져야 할 것이다.

참 고 문 헌

- 김선형. 1998. "과학기술정보 데이터베이스 품질평가에 관한 연구", 『서울여대 대학원 문헌정보학과 석사학위논문』
- 박준식. 1996. "색인의 평가기준에 관한 분석적 연구", 대구: 계명대학교.
- 안인자. 1997. "전자도서관의 WWW데이터 평가기준에 대한 연구", 『정보관리학회지』 1997, pp.249-67.
- 유사라. 1999. "데이터베이스 정보품질평가의 메타분석", 『정보관리학회지』, 16(1): 157-74.
- 유혜영. 1996. "국내제작 데이터베이스의 평가에 관한 연구", 『서울여대 대학원 문헌정보학 석사학위 논문』
- 이응봉. 1996. "데이터베이스 품질에 관한 비평적 평가", 『국회도서관보』, 33(4): 13-28.
- _____. 1999. "CD-ROM 데이터베이스의 품질 평가 및 개선방안", 한국문헌정보학회지, 33(4): 29-46.
- 이제환. 1997. "과학기술분야 서지 DB의 품질 관리 및 평가방안: KORDIC의 KRISTAL DB를 중심으로", 『한국문헌정보학회지』, 31(3): 109-34.
- _____. 1997. "분산체제로 구축된 통합 데이터베이스의 품질검증 및 개선을 위한 연구", 과학기술처
- 한국데이터베이스진흥센터. 1998. "데이터베이스 표준 DPCS 006-012 (1998)". 『서울: 한국데이터베이스진흥센터.』
- Armstrong, C. J., 1994. "CIQM Report on Database Quality. Quality: Deaf to Definition", Database, 17(6): 45-48.
- _____, 1994 "Databases and Quality: Why not try 'What You See Is What You Get?'", Managing Information 1, 28-30.
- _____, 1994. "Database Quality: How Users Can Help - And Be Helped!", Proceedings of the Eighth Annual Computers In Libraries International Conference, 105-109
- _____, 1994. "The Centre for Information Quality Management: 097-421 441 - a single 'phone number for all your woes!", Library Technology News, 12: 3-5.
- _____, 1996. The Quality of Publicly-Available Databases: WYSIWYG or What?. [cited 2000. 2. 26].
- Basch R., 1990. "Measuring the Quality of the Data: Report on the Fourth Annual SCOUG Retreat", Database Searcher, October 18-23.
- Brodie, M., 1980. "Data Quality in Information Systems." Information and Management 13(1): 1-12.
- Daniel, E., 1993. "Quality Control of Documents." Library Trends, 41(4): 64-664
- Dolan, D. R., 1992. "Quality control at the system level", Online, 16(2): 30-35
- Fox, C., 1994. "The Notion of Data and Its

- Quality Dimensions", *Information Processing & Management*, 30(1): 9-19.
- Granick, Lois. 1991. "Assuring the Quality of Information Dissemination: Responsibilities of Database Producers", *Information Services and Use*, 11: 117-136.
- Harry, V. & Oppenheim, C. 1993. "Evaluations of Electronic Databases: Criteria for Testing CD-ROM Products. Part 1 & 2." *Online & CD-ROM Review*, 17(4): 211-222 & 17(6): 339-368.
- Jacso, Peter. 1991. Coverage and Accessibility in Ulrich's Plus and EBSCO-CD. *The Serials Librarian*, 20(1): 1-35.
- _____. 1992. "CD-ROM Software, Dataware, and Hardware: Evaluation, Selection, and Installation". Englewood, Colorado: Libraries Unlimited, Inc.
- _____. 1995. Testing the Quality of CD-ROM Databases, in Basch, Reva. *Electronic Information Delivery*. Brookfield, VT: Gower, 141-68.
- _____. 1996. "Watching your Online Bottom Line". *Online*, 6, 20(4): 50-51.
- _____. 1997. "Content Evaluation of Databases". *Annual Review of Information Science and Technology*, 32: 231-267.
- Juntunen, Ritva., Mickos, Elisabet., Jalkanen, Tuulikki. 1995. Evaluating the Quality of Finnish Databases, in Basch, Reva. *Electronic Information Delivery*. Brookfield, VT: Gower, 205-219.
- Langlois, J., 1990. "CD-ROMs: Considerations before purchasing", *CD-ROM Librarian*, 5(11): 17-18.
- Lawrence, Barbara., Lenti, Tony. 1995. Application of TQM to the Continuous Improvement of Database Production, in Basch, Reva. *Electronic Information Delivery*. Brookfield, VT: Gower, 60-87.
- Medawar, K., 1995. "Quality in Library and Information Services: a Review", *Program*, 29(4): 379-94
- Oberts, Kristin K. 1995. Quality Assurance in the Information Service Environment, in Basch, Reva. *Electronic Information Delivery*. Brookfield, VT: Gower, 91-97.
- Quint, B., 1993. "The Database Connection: Full of Text", *Document Delivery World*, 9(3): 34-36.
- _____. 1995. Better Searching through Better Searchers, in Basch, Reva. *Electronic Information Delivery*. Brookfield, VT: Gower, 99-116.
- Richards, T. and Robinson, C., 1993. "Evaluating CD-ROM Software: A Model", *CD-ROM Professional*, 6(5): 92-101.
- Rowley J. & Slack, F. 1997. "The Evaluation of Interface Design on CDROMs." *Online & CDROM Re-*

- view 21(1): 3-14.
- Tenopir, Carol. 1990. Online Information Anxiety. *Library Journal*, 115(14): 119-139.
- _____. 1993. "Moving toward Quality". *Library Journal*, 118(10): 86-87.