

연재순서

- 1월호 국가지식정보통합검색시스템
<http://www.knowledge.go.kr>
- 2월호 국가문화유산종합정보서비스
<http://www.heritage.go.kr>
- 3월호 과학기술종합정보시스템
<http://science.knowledge.go.kr>
- 4월호 한국역사정보통합시스템
<http://kh2.koreanhistory.or.kr>
- 5월호 정보통신통합검색시스템
<http://www.itfind.or.kr>
- 6월호 문화예술종합정보시스템
<http://www.culture-arts.go.kr/index.jsp>
- 7월호 건설교통전자정보관
<http://www.codil.or.kr/main/index.html>
- 8월호 학술연구정보서비스
<http://www.riss4u.net/index.jsp>
- 9월호 해양수산학술연구정보검색서비스
<http://www.ocean.go.kr/>
- 10월호 정부출연연구기관지식정보검색시스템
<http://www.ikis.re.kr/jsp/main.jsp>
- 11월호 총평

국가지식정보통합검색시스템 (<http://www.knowledge.go.kr>)

한국데이터베이스진흥센터는 '지식정보자원관리사업'을 통해 구축된 지식정보 DB의 품질을 점검해 보기 위해, 매월 1종의 DB를 선정해 품질을 평가하고, 그 결과를 2005년도 'DB 서비스 평가' 기획 시리즈로 연재한다. 특히 이번 기획 평가는 사용자 관점에서 DB 품질을 조명해 보는 만큼, 지식정보 DB의 활용도 극대화를 위한 실천적 대안 마련에 도움이 될 수 있을 것이라 기대한다.

글_김선영 한국데이터베이스진흥센터 온라인사업팀장

우 리나라 IT 산업의 국가 경제 기여도가 경제개발협력기구(OECD) 회원국 가운데 가장 높다고 한다. 또한 초고속인터넷 보급률도 세계 최고 수준으로, 대한민국은 이제 명실상부한 IT 강국임에 틀림없다. 그러나 이와 같은 경쟁력을 유지하기 위해서는 더 많은 노력이 필요하다. 아직은 미약한 부분이 많기 때문이다. 특히 지식정보의 다양성과 활용이 미흡하다는 지적이 많은데, 이것은 우리나라 IT 산업의 약점이며 지식정보사회의 발전을 저해하고 있다.

이에 따라 정부는 2000년 1월 '지식정보자원관리법'을 제정해 국가의 주요 지식정보자원의 체계적인 관리와 보존을 제도화하고, '국가지식정보DB 구축사업 5개년 계획(2000년~2004년)'에 따라 '지식정보자원관리사업'을 추진하고 있다.

지식정보자원관리사업은 정부부처와 공공기관에 흩어져 있는 지식정보를 DB로 구축하고 상호 연계함으로써 국민들이 편리하게 이용할 수 있도록 한 국가적인 대규모 프로젝트로, 2000년부터 2004년까지 추

진된 1단계 사업에만 2,380억원이라는 막대한 예산이 투입됐다. 여기에 포함된 주제 분야만도 인문·철학·종교·사회·과학·기술·예술·언어·문학·정보통신·역사 등 10개 분야에 달하며, 주제별로 '국가문화유산종합정보서비스', '과학기술종합정보서비스', '한국역사정보통합시스템', '정보통신통합검색시스템', '문화예술종합정보시스템' 등을 구축해 서비스하고 있다.

특히 1단계 사업에서는 지식정보의 디지털화에 주력해 DB로 구축한 자료만도 2억1,000만 건에 이르며, 1998년 10.3%에 불과했던 지식정보의 디지털화율을 2004년에는 42.9%로 4.2배나 끌어올렸다고 한다.

이제 2005년을 맞아 '지식정보자원관리사업'은 제2단계 사업을 시작한다. 2단계 사업의 목표는 지식정보 DB의 활용도 극대화라고 한다.

그러나 DB의 활용성은 단순한 양적 확대만으로는 확보될 수 없다. 데이터의 양이 아무리 많아도 DB 품질이 열악하면 사용자들은 금방 외면해 버리기 때문이다.

국가대표 지식정보포털, '국가지식정보통합검색시스템'

지식정보 DB 기획 평가 첫 번째로, 지식정보자원관리사업의 핵심 결과물이라고 할 수 있는 '국가지식정보통합검색시스템'을 대상으로 선정했다.

국가지식정보통합검색시스템(이하 통합검색시스템)은 과학기술, 교육학술, 문화유산·예술, 역사, 정보통신, 건설기술, 산업경제, 해양수산 분야 등 전국 700여개 기관이 협력해 구축한 지식정보 DB를 윈스톱으로 검색할 수 있도록 한 우리나라 대표 지식정보 포털 서비스이다.

더구나 이용 방법도 매우 간단해, 통합검색, 필드검색, 상세검색 중 원하는 검색창에 키워드만 입력하면 10여개 분야의 지식정보 DB를 찾아 검색결과를 제공해 준다. 검색된 대부분의 자료는 원문까지 볼 수 있다.

특히 통합검색시스템에서 제공되는 자료와 DB는 분야별 전문기관에서 책임있게 만들어지고 있어, 데이터의 신뢰성은 그 어느 DB에도 뒤지지 않는다.

이처럼 통합검색시스템은 지식정보 사용자들의 정보 검색과 이용의 효율성을 획기적으로 개선한 서비스로 평가받고 있다.



▲ 국가지식정보통합검색시스템 메인화면

한국데이터베이스진흥센터 'DB 품질평가 모델' 활용

통합검색시스템 품질 평가는 한국데이터베이스진흥센터에서 개발한 'DB 품질평가 모델(ver.2.0)'에 기초해 수행했다.

DB 품질평가 모델은 데이터 값뿐만 아니라 데이터 구조, 데이터 표준 품질은 물론, DB 관리 프로세스와 시스템 품질까지 다양한 영역을 대상으로 평가를 실시함으로써 데이터 값에서 나타나는 현상(現狀)적 품질 문제의 원인을 밝혀 낼 수 있도록 고안된 종합적 평가모델이다.

그러나 본 DB서비스평가는 DB 운영기관의 동의 없이 한국데이터베이스진흥센터에서 대상 DB를 임의로 선정해 평가를 실시하고 있어, DB 구조나 운영 관리 프로세스 등은 진단할 수 없다는 제약이 있다. 따라서 본 평가에서는 데이터 값과 서비스 인터페이스와 관련된 품질 진단 항목을 위주로 평가 모델을 재구성해 사용했다. 평가에는 10명의 DB

전문가와 25명의 일반 이용자 등, 총 35명이 참여했다.

검색 성능 및 속도 개선 필요

데이터	완전성	★★★★☆☆	
	최신성	★★★★☆☆	
	정확성	★★★★☆☆	
	시스템	편의성	★★★★☆☆
		안정성	★★★☆☆☆
종합		★★★★☆☆	

▲ 국가지식정보통합검색시스템 평가결과

통합검색시스템 품질 평가결과, 5점 만점 기준에 전문가 평가 점수는 2.6점, 이용자 평가 점수 3.2점으로 총점 2.9점으로 나타났다. 항목별로는 데이터 완전성 2.9점, 데이터 최신성 2.6점, 데이터 정확성 3.2점, 그리고 시스템 편의성 2.9점, 시스템 안정성 2.4점으로, 통합검색시스템의 내재적 가치만으로 볼 때 매우 낮은 점수라 할 수 있다.

먼저 대부분의 평가자들은 통합검색시스템에서 제공하는 가치있는 방대한 양의 데이터를 높이 평가해, 데이터의 양과 범위, 신뢰성에 대해서는 높은 점수를 매겼다. 반면, 데이터의 최신성에 대해서는 보완이 필요하다는 의견이다.

데이터의 최신성은 사용자들의 DB의 활용 여부를 결정하는 중요한 판단 기준의 하나이다. 그러나 기존의 아날로그 형태의 자료를 디지털

〈표〉 DB 품질평가 모델의 평가 항목

평가영역	평가항목			
데이터	완전성	해당 서비스 목적에 적합한 데이터를 제공하고 있는가? 제공 데이터의 양과 범위는 서비스 목적에 충분한가? DB 항목(필드)은 해당 정보 특성에 적합하도록 구성되었는가? DB 항목(필드)에서 데이터 값이 누락된 항목은 없는가?		
	최신성	데이터는 지속적으로 현행화(update)되고 있는가?		
	정확성	데이터 값 또는 텍스트는 일관되게 표현되어 있는가? 중복 데이터 또는 철자 오류는 없는가? 제공 데이터는 신뢰할 만 한가?		
		서비스	편의성	제공 정보의 특성에 적합한 다양한 검색 기능을 제공하는가? 검색 결과는 활용 목적에 따라 원하는 형식으로 출력 가능한가? 검색 결과는 정보의 특성에 적합한 매체 혹은 형식으로 제공되는가? 메뉴, 페이지의 레이아웃, 배열이 일관성 있으며 보기 쉬운가? 페이지별 색상, 디자인 사용버튼, 아이콘 등이 적절히 사용되고 있는가? 사용자와 운영자간 커뮤니케이션 및 이용 안내가 적절한가? 회원 가입·탈퇴가 자유로우며, 개인정보보호 정책이 수립되어 있는가?
			안정성	전송속도 및 로딩속도는 만족스러운가? 사이트 이용시 페이지 또는 시스템 오류로 인한 문제는 없는가?



▲ 검색결과 오류

화 해 서비스하는 '지식정보자원관리사업'의 특수성을 감안할 때, 가장 최근의 데이터를 제공하라는 주문은 다소 무리일 수 있다. 그럼에도 불구하고 대다수의 전문가와 이용자들이 최신 데이터의 부족과 보완을 지적했다는 점은, 지식정보 DB가 활용성을 높이기 위해 앞으로 이에 필요한 대책을 세워야 함을 의미한다. 한편 메인화면의 구성과 디자인은 검색 서비스의 특성을 잘 반영하고 있다는 평가다.

DB 서비스의 핵심 기능이라 할 수 있는 검색창과 주제별 디렉토리가 눈에 가장 잘 띄는 위치에 황금률로 비중있게 배치돼 있어, 한눈에도 통합검색시스템이 전문 DB 서비스임을 알 수 있게 해 준다. 그러나 검색 결과 오류와 검색 속도에 대해서는 개선이 시급하다는 지적이다.

일반적으로 데이터 자체의 품질과 상관없이 검색시 발생하는 빈번한 오류나 검색속도의 저하는 사용자가 DB 검색을 쉽게 중단해 버리게 만든다. 그런데 통합검색시스템에서 검색된 자료의 상세화면을 보고자 할 때, '요청하신 문서를 찾을 수 없습니다' 라는 메시지가 뜨는 경우가 매우 많고, 간략화면 검색 자료의 제목 첫 글자가 절단돼 보이지 않는 등 크고 작은 문제들이 적잖이 발견된다. 또 주제와 전혀 상관없는 자료가 검색되기도 한다.

특히 검색속도도 느려 포털 사이트 검색에 익숙해 있는 사용자에게는 답답하게 느껴진다. 평가자들의 지적이 가장 많았던 상세화면 Dead Link는 특정 DB에서 유독 반복적으로 나타나는 경향이 있다. 이것은 간단한 DB 모니터링만으로도 쉽게 발견해 조치 가능한 비교적 단순한 오류로, DB 품질 확보에 있어 지속적인 관리가 얼마나 중요한지를 보여주는 단적인 예라 할 수 있다.

키워드 검색의 불완전성

이번에는 검색 기능의 품질을 좀더 자세히 살펴보기 위해 테스트 검색을 실시했다. 먼저 [인문-도서학-저작권법] 주제 분류에서 찾은 '디지털도서관과 저작권법'이라는 자료가 키워드로도 검색되는지 살펴보기 위해, '디지털도서관' 과 '저작권법' 을 키워드로 설정해 통합검색을

시도했다. 그러나 해당 자료가 검색되지 않았다.

또 [철학-논리학] 주제 분류에서 찾은 '삼단논법과 법논리' 라는 자료를 찾기 위해, 통합검색에서 '삼단논법' 과 '법논리', '삼단논법', '법논리' 등 세 가지 유형의 키워드 검색을 실시했다. 그러나 유일하게 '법논리' 를 키워드로 넣었을 때만 동일한 자료가 검색됐다.

본 평가에서는 통합검색시스템의 검색방식이나 검색엔진 성능 등에 대한 구체적인 정보가 없어 정확한 진단이 어려웠으나, 왜 검색 오류나 부실한 검색 결과가 나오는지 면밀하게 검토해 볼 필요가 있다. 그 밖에 검색결과를 최신 자료의 순이나 주제별 필터링을 위해 제공 기관별로 정렬할 수 없는 것도 아쉬웠다.

한편, 지식정보 DB별로 검색결과와 제공 형식이 달라 전체적으로 정제된 느낌을 주지 못하며 사용자에게는 혼란을 준다.

일반적으로 검색결과에서 제공되는 간략화면 정보는 자료의 적합성 판단 기준이 되며, 여기에는 제목, 저자, 제작년도, 제작기관, 초록 등이 포함된다. 그런데 통합검색시스템의 검색결과 간략화면은 지식정보 DB간에 일관성이 없을 뿐만 아니라, 자료에 따라서는 제목 밖에 제공되지 않은 경우도 많았다. 더구나 상세화면의 제공형식과 메타데이터는 DB간에 공통점을 찾기가 거의 힘들다.

예를 들어 '봉무정' 이라는 키워드로 검색을 하면 2건의 자료가 검색된다. 참고로 봉무정은 대구에 위치해 있는 조선시대 정자다. 검색결과는 각각 '문화유산정보시스템' 과 '문화예술종합시스템' 에서 제공되는 자료로, 시스템 별로 제공 형식은 완전히 다르나 제공 정보의 내용은 거의 유사하다. 차이가 있다면 '문화유산정보시스템' 에서는 '봉무정' 을 문화유산으로, '문화예술종합시스템' 에서는 작품으로 다루고 있는 정도다.

통합검색이라면 동일하거나 유사한 형식의 DB간에는 최소한의 메타데이터 표준화가 반드시 필요하다는 게 모든 평가자들의 공통된 의견이다.

우리나라가 IT 강국이 될 수 있었던 것은 정부의 적극적인 산업 육성 의지와 기업들의 투자, 그리고 국민 개개인의 높은 관심이라는 3박자가 잘 맞아 떨어졌기 때문이다.

통합검색시스템도 마찬가지다. 서비스의 활성화를 위해서는 정부와 기업의 노력은 물론 국민, 사용자의 관심을 집중시킬 수 있어야 한다. 그러나 2004년 11월 현재 통합검색시스템의 접속건수는 일평균 4,391건, 월 평균 13만1,715건이라고 한다. 국민들의 관심을 충분히 끌고 있다고 하기엔 아직 좀 부족한 수치다.

그런데 통합검색시스템에 사용자의 관심을 집중시키기 위해서는 그들의 요구수준(needs)에 맞는 품질의 서비스를 제공하는 것이 무엇보다 중요하다. 지식정보자원관리사업 추진 과정의 특성과 통합 검색에 대한 기술적 한계 등으로 어려움은 있겠지만, 금번 품질 평가를 통해 지적된 문제점들만 일부 보완된다면 충분히 가능한 일이 아닐까 생각한다. ●

한국데이터베이스진흥센터는 2004년 7월부터 매월 1종의 DB를 선정해 품질평가를 실시하고, 그 결과를 <디지털콘텐츠>의 'DB서비스평가' 코너와 데이터베이스품질평가센터 홈페이지를 통해 공개하고 있습니다. 자세한 내용은 홈페이지(<http://www.dbq.or.kr>)를 참조해 주십시오.

계제일	평가 DB	운영기관	평가점수
2004년 7월	Visit Korea	한국관광공사	★★★★☆
2004년 8월	서울시 버스노선 안내	서울특별시	★★★★☆
2004년 9월	농산물 유통정보 서비스	농수산물유통공사	★★★★☆
2004년 10월	워크넷	노동부	★★★★☆
2004년 11월	게임산업종합정보시스템	한국게임산업개발원	★★★★☆
2004년 12월	KIPRIS	한국특허정보원	★★★★☆