

캠브리지 과학 앱스트와 BIOSIS 비교

캠브리지 과학 앱스트랙트(Cambridge Scientific Abstracts, CSA)는 캠브리지 과학 앱스트랙트의 인터넷 데이터베이스 서비스(<http://www.csa.com>)의 제품과 제품 개선으로 최근 들어 바빠졌다. 이제 5.0 버전을 이용할 수 있다.

여기에는 확장기능이 첨가되었는데, 빠른 검색과 고급 검색을 필드 코드 선택과 부울 연산자 선택을 할 수 있는 드롭다운 박스로 단순화시키고, 복잡한 부울 검색을 구성하는데 익숙한 사람들을 위해 커맨드 라인 박스를 제공하며, 아카이브와 올해를 단일 검색으로 결합하고 10단어 이내로는 근접 검색이 가능하도록 하고 있다.

이 제품에 대해 아이다호주가 처음에 가졌던 관심은 생물과학의 주요 도서목록 데이터베이스에 데스크탑 전자 액세스를 제공하려는 바람에 기초하고 있었다. 그리하여 그 제품을 온사이트 상태로 시험가동하고 개괄적으로 살펴보기로 결정되었다.

이 시험기간 동안(이때는 5.0베타버전이었음), CSA 데이터베이스가 훌륭한 기능을 가졌고, 가격이 적당하며, 응답 시간이 빠르게 대륙 서부의 고객들을 위해 별도로 서버를 제공하는 것과 같은 세심한 주의를 기울였다는 것을 알게 되었다. 이 시험으로 CSA의 인터넷 데이터베이스 서비스는 우리 스스로 시험을 정당화할 만큼 충분히 매력적이라는 것이 드러났다. 이 기간 동안 이 제품의 콘텐츠와 다른 두개의 콘텐츠를 비교하였다. 그 두 가지는 Basic BIOSIS와 BIOSIS Previews이다.

시험기간을 30일로 정하고, CSA의 생물과학 수집자료 데이터베이스에 특히 초점을 맞췄다. 이 데이터베이스는 데스크탑형이고, 인터넷에 기반을 두고 있으며, 방대한 수의 동시 사용자 사이트 라이선스를 받은 제품으로, 2백 만개 이상의 레코드를 보유하고

고 있다.

이 수집자료에는 Elsevier의 식물과학 색인, Plexus Publishing의 생물학 다이제스트와 캠브리지 전유의 생물과학 데이터베이스가 포함된다. 이 시험기간 동안 CSA 데이터베이스와 DialogWeb CIP의 BIOSIS Preview와 OCLC'S FirstSearch의 Basic BIOSIS를 비교해보기로 계획했다.

DialogWeb CIP과 BIOSIS Preview의 딜리버리

DialogWeb CIP(교실 교육 프로그램)에 관한 한 마디: 오블리 도서관 참고도서관 부서에서는 아이다호 주립대학교 학생들에게 오랫동안 CIP를 지원했고, 여러 차례 적극적으로 판매했었다.

현재 Dialog의 CIP의 가격은 우리 대학에서 제공하는 액세스를 전부 보조할 수 있는 정도이기 때문에, 우리의 Dialog 워크숍에 참가하는 학생들에게 학과 과정 관련 과제와 논문 관련 연구과제를 하도록 대다수의 Dialog 데이터베이스에 액세스할 수 있도록 하고 있다.

Dialog의 조건에 의해 아이다호 주립대학에서는 Dialog를 정규 도서목록 교육 프로그램과 관련하여 가르친다. 대부분의 상급 학생들과 대학원생들은 BIOSIS Preview와 같은 자연과학 데이터베이스의 자료를 얻기 위해 CIP 액세스를 요구한다. 그렇지 않았다면 지금도 이들 학문 분야에서 색인하고 발췌하기 위해 종이문서에 의존하고 있을 것이다.

CSA의 생물 과학 수집자료 데이터베이스와의 비교에서 흥미 있는 것에는 으뜸가는 생물 과학 도서목록 데이터베이스인 BIOSIS Preview에 제공된 CIP 액세스가 있었다. 교직원들은 CIP를 이용할 수 없다는 것이 결점들 중 하나이다(적어도 교직원에게

게는 그렇다). 그러나 그 조항에 따른 교육 규정에도 불구하고 우리에게 그것은 계약을 통해 거래할 수 있는 여지가 있는 서비스이다.

OCLC의 FIRSTSEARCH를 경유하는 BASIC BIOSIS

OCLC의 FirstSearch는 다년간 학술 도서관 내에서 온라인 정보를 제공해온 정보제공업체로서 현재 대대적으로 기능을 확장하고 있다. 새로운 FirstSearch의 출시는 1999년 8월 하순에 시작되어, 그 해 말에 모든 기능을 완전히 실행하게 되었다.

이들 강화된 기능 중에 가장 두드러진 것은 데이터베이스 교차(cross-database) 검색, 절단(장시간의 지연), OR 부울 연산자의 무제한 사용이 포함된다. 또한 어구 검색에 인용부호를 이용할 수 있게 될 것이고, 그로 인해 문맥 감지 도움말, 관련 등급, 데이터베이스 시소러스 수의 증가가 가능해질 것이다.

우리는 최초의 FirstSearch의 블록 검색 액세스를 이용하여 Basic BIOSIS 데이터베이스를 시험했다. OCLC에 맞춰 특수하게 고안된 Basic BIOSIS 데이터베이스에는 OCLC의 조합 카탈로그에 제시된 것처럼 가장 일반적으로 보유한 생물과학 저널들에 대한 개요들과 더불어 인용문들이 들어있다.

이 외에도 신문 수집자료와 Audubon, Natural History와 같은 제목의 일반적 관심사를 다룬 잡지들도 들어있다. 여기서는 1994년의 자료부터 400,000만개 이상의 레코드가 포함된 데이터베이스를 제공한다. BIOSIS는 연구 파일인 보다 규모가 큰 BIOSIS 파일과 같지 않게 대학생들을 이 데이터베이스의 표적으로 삼고 있다.

이런 비교가 전적으로 공정한 것은 아니라는 것을 인식하고 있다. 왜냐하면 BIOSIS 데이터베이스의 전체 규모는 CSA의 생물과학 수집자료 데이터베이스보다 5배 더 크고, FirstSearch의 Basic BIOSIS 데이터베이스보다 20배나 더 크기 때문이다. 본질적으로 CSA 데이터베이스가 BIOSIS 데이터베이스를 대체하지는 못한다는 것을 솔직히 인정한다. 적어도 포괄적인 검색을 요하는 학생들에게는 그렇지 못하다.

우리가 진정으로 하고자 했던 것은 CSA 데이터베이스를 평가하는 것이었는데, 그 방법은 CSA를 콘텐츠의 목적과 가격 범위에 입각해서 논리상의 경쟁자들과 비교하고, 재무 위원회에 회부했을 때 생물과학 전자 출판물 데이터베이스를 고려하는 도서관

들이 돈을 잘 썼다고 판단하는지에 의한 것이었다. Basic BIOSIS 데이터베이스와 BIOSIS Preview 데이터베이스 모두는 우리에게 비교를 위한 벤치마크를 제공하게 될 것이다.

리트머스 시험

우리는 각 데이터베이스에서 똑같은 검색을 42번씩 실행했다. 각 검색엔진의 매개변수들에 기초하여 검색어들을 세심하게 조합했다. 몇 가지 경우에서 처음의 FirstSearch 인터페이스에 의해 제한을 받았다. 그로 인해 절단 기능을 사용하지 못하거나, 매 검색에서 OR 부울 연산자를 세번 이상 사용할 수 없었다.

우리는 나중에 새로운 FirstSearch 플랫폼에서 그들 중 샘플을 택해 검색을 해봄으로써 Basic BIOSIS 데이터베이스의 콘텐츠가 별로 변하지 않아서 확장된 검색 기능들이 검색자료를 눈에 띄게 증가시키지는 못했다는 것을 확인했다.

검색 결과의 질에 대한 평가는 하지 않았다. 우리가 다룬 것은 출력된 그대로의 검색자료였고, 정확성보다는 호출에 초점을 맞췄다. 우리는 정말 가장 복잡한 것으로 평가된 여덟 가지 검색결과를 그대로 내어놓았다(그들에는 세 가지 명확한 개념이 있음).

BIOSIS Preview 데이터베이스가 가장 인상적이었는데, 42번 이상의 매 검색에서 평균적으로 200개 이상의 결과자료를 내어 놓았고, 42번의 검색을 통해 얻은 전체 자료의 67퍼센트 이상을 차지했다. BIOSIS 데이터베이스는 복잡 검색으로 검색의 한계를 지원했는데, 그 비율은 대개 3대 1이었다.

생물과학 데이터베이스와 생활하기

우리는 캠브리지가 고급 검색 기능을 갖춘 훌륭하고 안정된 인터페이스를 가졌긴 하지만, 그 콘텐츠에는 결점이 있다는 결론에 도달했다. 특히 대학원에 재학 중인 생물학도들과 교수들에게는 그랬다.

훌륭한 검색 인터페이스와 중간 정도의 가격, 곧바로 신원확인 난을 채울 필요가 없음에도 불구하고 CSA의 생물과학 수집 자료 데이터베이스를 떠난 대학원생들에게 Basic BIOSIS 데이터베이스와 윌슨의 생물학/농학 색인 데이터베이스는 상당히 도움을 줄 수 있다.

이것은 장애가 없고 직관적인 검색 인터페이스의 질을 알봐도 된다는 뜻은 아니다. 그리고 이것은 캠브리지가 제공하는 다른 많은 매력적인 데이터베이스 수집 자료들에 대한 관심을 박탈하

는 것 또한 아니다.

그러나 생물학 수집 자료 데이터베이스가 BIOSIS 데이터베이스와 견줄 수 있다는 의미를 함축하고 있는 것이라면(필자는 CSA의 메시지에 그런 의미가 담겼다고 생각한다), CSA는 그를 위해 데이터베이스의 콘텐츠를 증가시키는데 초점을 맞춰야만 한다.

우리는 CSA의 생물과학 수집 자료 데이터베이스에서 "현재 포함된 것으로 모니터 된" 3,696개의 정기간행물 중, 716개, 즉 19 퍼센트만이 캠브리지에 의해 책의 전권이 색인되었음을 나타내는 "핵심 소스들"로 간주되는데 관심을 가졌다. 캠브리지의 데이터베이스를 비용 효율적이라고 인식하는 사람들도 있었지만(제한 없는 사용자 사이트 라이선스를 얻는데 9,995달러), 최종 분석에서 많은 도서관들은 BIOSIS Preview 데이터베이스와 Dialog를 통해 액세스할 수 있는 다른 생물과학 데이터베이스에 액세스하기 위해 Dialog CIP 프로그램을 구입하고, 대학생들을 대상으로 한 FirstSearch의 Basic BIOSIS를 제공한 것에 보다 더 안도감을 느끼게 될 것이다.

그런 사이에 그들은 한숨을 돌리면서 캠퍼스의 데스크탑에 Basic BIOSIS 데이터베이스를 배달하라는 요구에 대해 결단을 내리고자 애를 쓴다.

마지막으로 캠브리지의 인터넷 데이터베이스 서비스의 5.0버전이나, 근접검색 확장 기능, 키워드검색 탄력성 확장 기능, 보다 깔끔해진 기본/고급 검색 스크린, 기타 개선된 기능들로도 그것의 생물과학 수집자료 데이터베이스와 BIOSIS 데이터베이스 사이에 존재하는 콘텐츠의 격차를 메울 수는 없었다.

■ 저자/ 제임스 자트케비시우스의 전자우편 주소는 jatkjim@isu.edu이다.

캠브리지 과학 애브스트랙트에서 보내온 대답

제임스 P. 머킨티 사장

자트케비시우스씨는 자신의 기사에서 몇 가지 논지를 아주 훌륭하게 입증하고 하고 있다. 캠브리지 과학 애브스트랙트(CSA) 인터넷 데이터베이스 서비스(IDS)의 현 5.0버전의 기능과 새롭게 확장된 기능들에 대한 그의 비평은 고마운 것이었다. 그에 더해 도서관의 가용 예산과 특정 고객에 대한 서비스 사이의 균형을 취하는 일에 관한 그의 견해는 모든 도서관들에게 있어 중요한 문제이다.

그렇긴 하지만, 그의 기사로 인해 생길 수 있는 몇 가지 오해를 바로잡고자 한다. 캠브리지 과학 애브스트랙트 데이터베이스는

42번의 검색을 통해 얻은 총 검색 결과

질문 BIOSIS Preview Basic BIOSIS CSA

가금류에서 식중독 박테리아의 검출 129 6 67
 신경계의 발달단계에서
 FGF나 bFGF의 영향 * 1280 291 108
 거미의 식습관 46 7 22
 Sphaerotilus natans의
 살균바이러스 박테리오파지 2 2 1
 담배식물에서 꽃가루 생성 유전자 조절 * 186 68 132
 초식동물과 군서의 역학 * 155 45 195
 식물 뿌리 부패에 있어 이산화탄소의 영향 4 5 6
 인간의 vomeronaal organ의 기능 * 20 4 11
 흙탕물에서 땅 다람쥐의 총격 6 0 4
 physarumpoly cephalum에서
 aminine의 신진대사 15 2 4
 단백질 합성에서 ACTH의 영향 217 4 25
 식용우 가축 사육장에서 밀파리 폐에
 미치는 살충제의 효과 * 4 2 2
 누룩병의 체온 조절 18 3 13
 무지개 송어의 중금속 축적의 생리학적 영향 7 1 6
 토양 침식으로 인한 기후 변화의 영향 15 0 11
 트리클로로에틸렌에 의한 지하수 대수층 오염 311 47 125
 감자의 가뭄 내구력 24 0 8
 아메리카 들소와 부르셀라증 82 25 27
 Lodgepole pine와 딱정벌레 163 22 113
 늑대와 재도입 32 44 25
 고가집이와 대서양 대구 178 30 783
 방목지와 공유지 26 21 14
 큰사슴과 기생충 64 6 35
 야생백조와 서식지 9 2 15
 Blatomyces dermatitidis균 996 12 308
 방귀벌레와 거미 10 0 26
 포유동물에서 약물로 인한 기형 2353 86 31
 물새와 대형 무척추 동물 67 9 33
 미생물에 의한 생물분해와
 테트라클로로에틸렌 또는 트리클로로에틸렌 255 73 255
 고생물학과 아이디호주, 오레곤주, 혹은 워싱턴주 248 0 2
 bitterbrush 새맛과 번식 9 3 6
 펭귄과 먹이약탈 252 52 160
 페루 앵무새 12 0 3
 버과의 풀 78 6 16
 얼룩무의 황어의 서식지 27 6 26
 방울뱀과 포식성 57 11 11
 도마뱀과 운동 276 66 119
 Tiger mosquitoes와 땀그늘병 287 10 123
 질소와 토양과 사막의 식물* 286 21 116
 에틸렌 다이브로마이드와 잔류물 61 2 18
 코요테의 생존 65 16 42
 콜로라도 감자 딱정벌레의 살충제 내성 126 5 75

* 별표는 여덟 가지의 가장 복잡한 검색을 나타낸다.

CSA 생물과학 데이터베이스가 BIOSIS Previews의 맞적수가 된다고 주장한 적이 결코 없었다. IDS를 제공하고 있는 CSA 생물

과학 데이터베이스(<http://www.csa.com/details/biolset.html>)는 BIOSIS Previews 데이터베이스를 구입할 여유가 없는 단체들에게 비용 효율적인 대안으로서 자리하고 있다.

그러나 CSA 데이터베이스는 FirstSearch의 Basic BIOSIS 데이터베이스와 윌슨의 생물학 색인 데이터베이스 양자나 혹은 둘 중 어느 하나에게 있어서는 명백한 경쟁 상대이다. 자트케비시우스씨도 CSA 데이터베이스가 콘텐츠 범위에 있어 FirstSearch의 Basic BIOSIS 데이터베이스보다 월등하다고 말한 것으로 보아 그도 이점에 동의하는 것으로 보인다.

또한 필자는 인터넷 데이터베이스 서비스(ISD)의 가치가 낮게 평가되었다고 여기고 있다. 9,995 달러의 가입비만 내면(3,500 FTE보다 수가 적은 학교들에는 6,600 달러) CSA 생물과학 데이터베이스 사이트 전체에 무제한 액세스할 수 있다. 이 데이터베이스에는 22개의 서브 파일이 있고, 그 파일들은 2백 2십 만개 이상의 레코드로 구성하고 있다.

그리고 그것들은 곤충학과 동물 습성에서부터 생물 공학과 신경 과학에 이르기까지 광범위한 주제들을 다루고 있다. 또한 정기가입 서비스에 포함되는 것에는 생물학 다이제스트, 회의 문서 색인, MEDLINE(10년간), 식물 과학, TOXLINE(5년간), 인터넷 리소스 데이터베이스, 전체 6백 5십 만개의 레코드의 최근의 참고도서들을 포함하여 여타 생물학 관련 데이터베이스들의 액세스가 있다. 그것도 하나의 커맨드로 전부 검색할 수 있다.

필자로서는 42번 실행된 검색에 대한 구체적인 검색 전략을 알지 못하지만, 검색비율의 격차에 어안이 병병하고 그에 대한 나름대로의 생각을 가지고 있다. Dialog의 BIOSIS Previews 데이터베이스에는 1969년부터 지금까지의 자료가 들어있다. 그리고 인터넷 데이터베이스 서비스 파일들은 다양하지만, 주로 1982년부터 현재까지의 자료들을 다룬다.

이것만으로도 검색된 참고자료의 수는 상당히 달라질 수 있다. 그러나 검색자료들의 숫자는 그 자체로서 서비스를 선택하기 위한 훌륭한 표준은 되지 못한다. 관련성이 중요하다. 자트케비시우스씨도 인정했듯이 관련성을 평가하려는 시도는 전혀 이루어지지 않았다. 가격 또한 요인이다. 그가 쓴 기사에서는 CSA의 가격에 대해서는 언급했지만, Dialog CIP 프로그램의 비용에 대한 언급은 없다. FirstSearch의 가격 구조에 대한 언급도 전혀 없다. 예를 들어 CSA 생물과학 데이터베이스가 제공하는 인터넷 무제한

한 액세스 서비스의 가격은 생물 애브스트랙트의 인쇄본을 일년간 정기 구독하는 가격과 엇비슷하다. 경제적 준거가 없다면, 비교가 불가능해진다.


자트케비시우스씨의 비교에서 가장 중요한 측면은 라이선스에 의해서 CIP 프로그램을 아이다호 주립 대학의 교직원들은 이용할 수 없다는 사실에 있는 듯하다. 이것은 다른 요인이 이 방식식에 부가되어야 함을 뜻한다. 다시 말해 연구직 교직원들에게 그 데이터베이스로의 액세스를 제공하는데 드는 비용이 부가되어야 한다.

CIP 프로그램과 같은 서비스를 캠퍼스의 모든 데스크탑에서 이용할 수 있는 무제한 액세스 서비스와 비교해도 되는 것일까? 그 프로그램은 교직원들은 이용할 수 없고, 도서관 1층 참고도서 영역에서만 학생들이 액세스할 수 있는 프로그램이다.

그리고 인터넷 데이터베이스 서비스에서 이용할 수 있는 독특한 콘텐츠도 간과되었다. 예를 들어, 인터넷 리소스 데이터베이스(<http://www.csa.com/details/ird.html>)는 액티브 URLs을 가진 2만 6천개의 생물학과 관련 색인 레코드들이 그 특징을 이루고 있다.

또한 최근 참고자료 데이터베이스(<http://www.csa.com/details/britlib.html>)는 매일 업데이트되며 1천개 이상의 생물과학 저널들에서 가져온 4만 5천개 이상의 최근 기사들이 들어 있고, 도서관 보유자료들로의 링크와 문서 딜리버리로의 링크도 제공하고 있다. 기타 다른 기능들에는 도서관 보유자료와 문서 딜리버리로의 링크뿐 아니라 검색 결과에서 직접 4천개 이상의 전자 저널 전문 기사들로의 링크도 있다. Dialog의 교실 교육 프로그램(CIP)에는 이러한 장점들에 필적할 만한 것이 없다.

CSA 데이터베이스의 우리들로서는 제품 비교를 환영한다. 그것은 도서관원들이 교육적 요구를 충족하는 최상의 서비스를 선택하는데 있어 그것들이 중대한 도구가 된다는 것을 인식하고 있기 때문이다. 우리는 검색자료의 양에 더해 검색자료의 관련성을 평가하고, 중복된 저널이 얼마나 되는지를 검토할 수 있는 그런 비교가 행해지기를 바란다.

우리는 이러한 콘텐츠 표준에도 불구하고, 재정적 요소가 참된 비교를 위해 중요하다고 생각한다. 자트케비시우스씨가 자신의 비교연구에 가격요소를 포함시켰었다면, CSA 생물과학 데이터베이스가 그곳의 학생들과 교직원들이 가장 잘 액세스 할 수 있는 비용 효율적인 대안이라는 것을 알게되었을 것으로 생각한다. 

■ 제임스 자트케비시우스/ 아이다호 주립대학 엘리 M. 오블러 도서관