

비즈니스맨을 위한 전문 정보검색 서비스

인터넷을 통해 정보검색을 할 수 있는 방법은 크게 검색엔진을 이용한 방법과 전문 데이터베이스를 이용한 방법을 들 수 있다. 이 가운데 검색엔진을 이용하는 경우는 사용자들에게 많이 소개되어 있지만 실질적으로 고급 정보가 빈약하다. 그러나 전문 데이터베이스(90%이상 무료)를 이용하는 경우 정형화되고, 양질의 정보를 구할 수 있다. 이 글에서는 인터넷을 통한 온라인 시장조사, 비즈니스 정보검색을 제공하는 서치플러스에 대해 살펴보고자 한다.

- 김화수/(주)칼스테크 기획개발실장
(topmost@calstec.com)

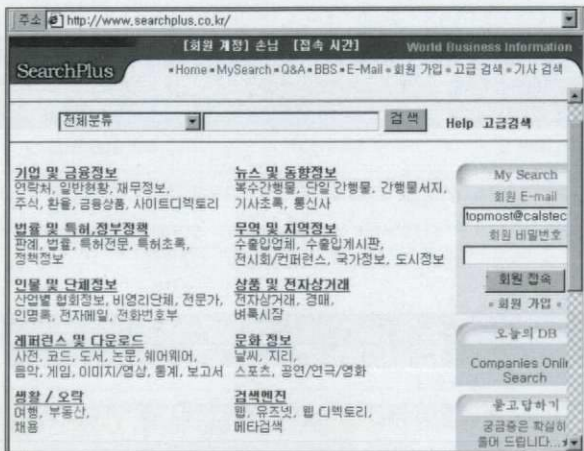
1. 인터넷 정보 공간과 전문 데이터베이스의 가치

인터넷의 정보 공간은 크게 웹과 유즈넷, 그리고 전문 데이터베이스로 접근 가능한 콘텐츠로 구분할 수 있지만 지금껏 인터넷 정보 검색을 논할 때에는 항상 검색엔진이 그 대상으로 검토되어 왔다.

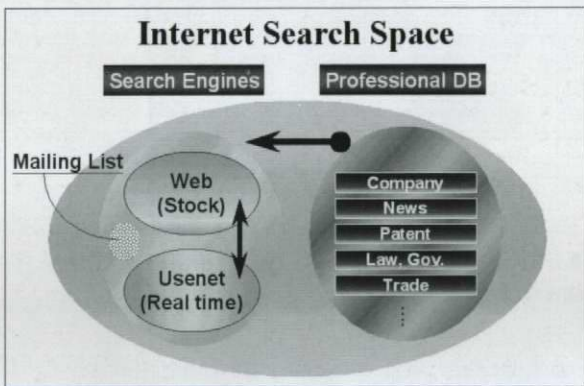
그러나 흔히 '검색엔진'으로 일컫는 AltaVista, Yahoo!, Deja와 같은 디렉토리, 전문(fulltext) 색인 시스템들은 에이전트의 접근이 허용되는 일반적인 웹 문서나 유즈넷 기사만을 그 검색 대상으로 하므로 '전문DB'에서 검색 가능한 콘텐츠에 대한 접근이 극히 제한적이다.

그러면, 이렇게 자주 접하는 검색엔진을 이용하여 검색 가능한 영역이 인터넷 전체 정보자원에서 차지하는 비중이 어느 정도 되는지를 파악해 보는 것도 흥미있는 주제가 될 수 있다.

검색엔진은 크게 그 검색 공간에 따라 웹 검색엔진과 유즈넷 검색엔진으로 나눌 수 있으며, 전문 데이터베이스들은 흔히 다루는 주제가 상이하므로 주제에 따라 기업정보, 법률정보, 무역정보, 특허정보, 산업별 정보 등으로 구분할 수 있다. 이러한 구분을 그림으로 표현하면 다음과 같다.



(그림 1) 서치플러스 메인 화면(<http://www.searchplus.co.kr>)



(그림 2) 인터넷 정보 공간 및 검색도구

물론 웹, 유즈넷, 전문DB로 구분했지만 각 정보공간에 따라 단위 정보의 규모 차이로 수치상의 비교에 어느 정도 문제가 있지만, 이러한 제약 요인에도 불구하고 전체적인 규모를 가늠하고 검색하는 것과 그렇지 않은 것은 숲을 볼 수 있는 능력을 지니는데 있어, 또는 정보검색 능력에 중요한 DB선택 능력을 높이는데 중요하다고 볼 수 있다.

인터넷의 규모는 어느 누구도 정확히 예측하는 것이 쉽지 않다. 그러나 몇 가지 방법을 통해 대략적인 규모나 상대적인 비교는 가능하다. 실제 NEC Research Institute와 같은 연구소는 정기적으로 이러한 통계를 조사하기도 한다. 이 통계에 대한 상세한 자료는 아래 URL에서 확인할 수 있다.

How big is the Web? How much of the Web do the search engines index? How up to date are the search engines?

<http://www.neci.nj.nec.com/homepages/lawrence/websize.html>

이 통계의 여러 조사들을 확인해 보면 1999년 9월 현재 웹의 규모는 약 50억 페이지, 유즈넷의 규모는 최근 1년간 포스

팅된 기사를 기준으로 계산할 때 약 5억개의 기사가 검색 가능한 수준임을 파악할 수 있다. 그렇다면 인터넷으로 접속 가능한 (무료로 제공되는) 전문DB에서 검색 가능한 자료는 어느 정도 규모일까? 전문DB의 규모는 100억~200억개 정도의 레코드가 인터넷으로 검색 가능한 규모라 할 수 있다.

실제 인터넷을 검색하는 많은 사용자들이 웹과 유즈넷 정보를 검색할 수 있는 일반 검색엔진으로 검색 가능한 정보원보다 인터넷에는 검색엔진으로 검색할 수 없는 자료들이 전문DB속에 존재하고 있다는 사실을 간과해왔다고 볼 수 있다. 또한, 단순 수치 비교로 볼 경우 50:5:150 정도로 볼 수 있지만 데이터의 신뢰성과 정형성은 웹보다 유즈넷이, 유즈넷보다 전문DB가 훨씬 높다고 보면, 그 만큼 전문 데이터베이스의 활용이 인터넷 검색에 있어 검색 가능성과 효율성을 증가시키는데 중요하다고 볼 수 있을 것이다.

이는 인터넷 정보검색 능력에도

영향을 미친다고 볼 수

있는데, 상호간에

중복 데이터가

많고, 따라서

전세계 수천

개의 검색엔

진 가운데

Best 10 정도

만 효과적으로 활

용함으로써 웹과 유

즈넷의 대부분 자료를 검

색할 수 있는 능력을 'Know How'의

영역이라고 본다면(일부 검색엔진에 대한 정확한 검색 방법만 파악하면 되므로), 전문DB 검색은 검색방법은 쉽지만(명확한 필드 검색을 지원하므로) DB간 데이터 중복성이 매우 낮아 최대한 다수의 데이터베이스를 검색할 수 있어야 한다. 때문에 이는 경험과 시간이라는 변수가 '검색능력'을 높이는데 중요한 변수가 될 수 있으며, 결국 'Know Where'의 문제로 귀결된다고 볼 수 있다.

또한 비즈니스, 과학기술 정보를 제공하는 전문 데이터베이스 검색은 아직 기술적, 법적인 문제로 인해 동시에 수백~수천 데이터베이스를 검색하는 것은 어렵다. 이는 검색엔진을 통해 수십만 웹사이트에 담겨진 문서를 동시에 검색할 수 있



〈표 1〉 인터넷 검색 도구 비교

분류	자료 성격		대표적 검색 도구			
	현재시점 자료	과거 -현재	디렉토리		전문(fulltext) 검색	
			국내	외국	국내	외국
웹	●	-	yahoo.co.kr	yahoo.com	naver.com	altavista.com
유즈넷	-	●	news.netple.com	tile.net	usenet.hanmir.com	deja.com
전문DB	●	●	searchplus.co.kr	search.com	-	-

〈표 2〉 일반 검색엔진/전문 데이터베이스 비교 (1999. 09. 기준)

비교항목	구분	일반검색엔진	인터넷 전문DB
서비스 수		2000개 이상	100,000개 이상
DB별 보유 자료 수		100,000,000 이상	10,000-50,000,000
색인		프로그램(Robot agent)	사람에 의한 수작업
검색공간		인터넷 내부 문서 검색	인터넷 외부 문서 연동
데이터 정형성*		비정형적	정형적
검색 기능		다양	단순
로봇이용		이용	이용하지 않음
상대적 검색 난이도		어려움	쉬움
요금		무료	유료/무료 구분, 유料的 경우 대부분 정액제
예		AltaVista, Infoseek, Hanmir 등	EDGAR, QPAT, NewsPage 등

* 데이터의 정형성이 높다는 것은 해당 데이터베이스가 보유하고 있는 모든 레코드들의 필드 구조가 동일하고 Null data(비어있는 값)가 없음을 의미합니다. 따라서 많은 사람들이 작성한, 개방적인 문법을 갖는 일반 HTML 문서를 대상으로 색인하는 검색엔진의 정형성은 당연히 낮을 수밖에 없습니다. 이는 정밀한 검색이 어렵다는 것을 의미합니다. 대부분 검색엔진들이 이를 해결하기 위해 다양한 검색기법과 독특한 검색결과 배열 방법들을 새롭게 개발하고 있는 것입니다.

는 것과 큰 차이가 있기 때문이다.

결국 전문 데이터베이스는 디렉토리 서비스를 통해 특정 분야의, 특정 지역의, 특정 산업의 정보를 제공하는 데이터베이스를 찾아갈 수 있는 '위치정보' 서비스가 필요하다. 따라서 본고에서는 이러한 전문DB 디렉토리 서비스를 제공하는 서치플러스(searchplus.co.kr)에 대해 기능과 검색 범주에 대해 살펴보기로 한다. 참고로 각 정보공간별, 자료 성격별, 검색형태별로 주요 인터넷 검색 도구들을 비교하고 대표적인 서비스를 소개하면 〈표 1〉과 같다.

2. 검색엔진과 전문 데이터베이스 비교분석

이제, 보다 구체적으로 일반 검색엔진과 비즈니스, 기업정보를 제공하는 인터넷 전문 DB를 비교해 보도록 한다. 9가지 항목을 기준으로 두 데이터베이스를 비교하면 〈표 2〉와 같다.

웹과 유즈넷 공간을 검색하는 경우 다소 까다롭긴 하지만 어느 정도의 시간을 투입하면 충분히 그 사용법을 익힐 수 있으며, 숙지해야 하는 검색엔진도 10개 안팎이면 충분하다. 그러나 전문 DB를 이용한 검색은 우선 데이터베이스가 위치하고 있는 소재(URL) 파악이 쉽지 않기 때문에 검색에 많은 어려

움이 따른다. 이러한 데이터베이스에 담겨진 자료들은 게이트웨이를 통해 검색결과를 제공하며 기술적, 법적으로 로봇 에이전트의 접근을 허용하고 있지 않으므로 콘텐츠 전문(fulltext)에 대한 색인이 사실상 불가능하다.

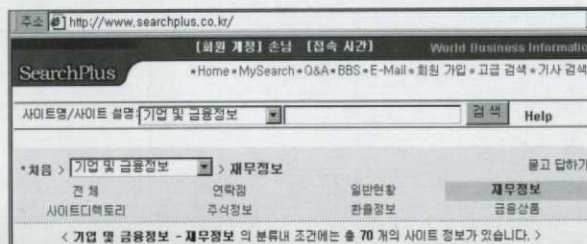
결국 해당 데이터베이스로 접근할 수 있는 검색 모드가 제공되는 곳까지 사용자가 직접 이동하지 않으면 검색이 불가능하다는 것인데, 문제는 일반 사용자가 그 곳이 어디인지 모른다는 것이다. 따라서 전문 데이터베이스 검색은 'Know Where' 능력을 요구하는 것이다. 'Know Where'의 필요성은 결국 디렉토리 서비스의 필요성을 말해준다.

3. 비즈니스 전문DB 디렉토리 서치플러스 (searchplus.co.kr)

서치플러스는 대분류 10개, 소분류 50여개로 구성된 전문 데이터베이스 디렉토리 서비스로 분류별 검색과 키워드 검색을 모두 지원하고 있다. 분류별 검색은 분류명 선택으로 최종 결과로 바로 이동할 수 있는 기능이며, 키워드 검색엔진 데이터베이스명이나 요약 설명 또는 데이터베이스 주소(URL)를 통해 검색, 결과로 이동할 수 있는 서비스를 제공한다.

먼저 분류 검색부터 살펴보기로 하자. 분류 가운데 '기업 및 금융정보 => 재무정보'를 선택하여 60개 데이터베이스가 등록되어 있음을 확인할 수 있다. 즉 이 분류는 기업 재무정보를 제공하는 전세계 데이터베이스 가운데 searchplus.co.kr에서는 70개 비즈니스 정보 데이터베이스에 대해 설명하고 있음을 알 수 있다.

1차 분류 및 2차 분류에 언제든지 이동할 수 있도록 네비게이션을 지원하고 있다. 여러 기준의 색인을 통해 사이트명 알파벳순, 분류별 보기, 방문 포인트별 보기 등 다양한 사용자 관점을 적용하고 있다.

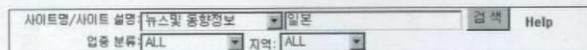


〈그림 3〉 검색 시작 화면



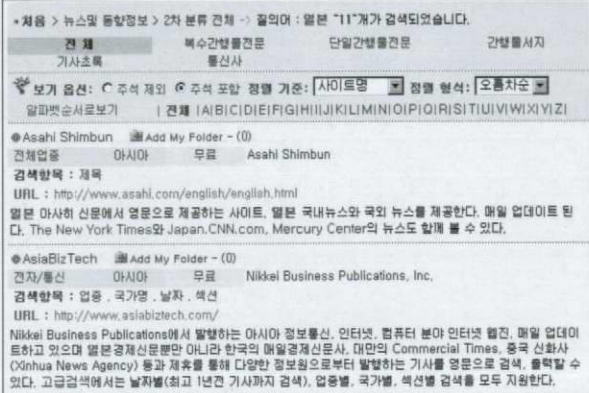
〈그림 4〉 검색 결과 화면

한편, 단순 키워드 검색의 경우 해당 키워드가 요약 내용 중 검색결과로 도출될 경우 해당 키워드에 색상을 달리해서 표현한다. 전체 검색과 분류내 검색을 동시에 제공하므로 쉽게 활용할 수 있을 것이다. 단순 키워드 검색 화면 및 검색결과 모습은 〈그림 5〉, 〈그림 6〉과 같다.

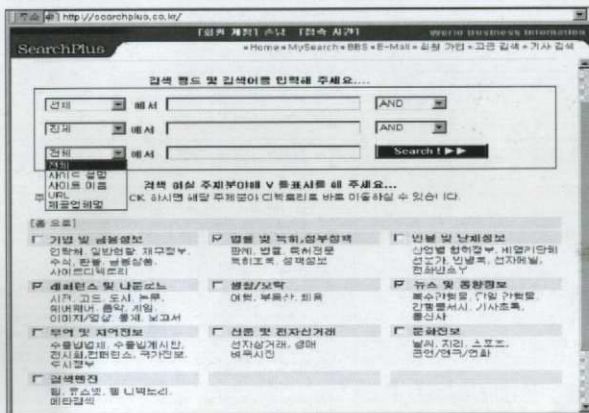


〈그림 5〉 키워드 검색 화면

단순 검색과 달리 고급(조건)검색은 다양한 연산자와 필드 제한 검색을 지원하는 조건 검색 모드를 별도로 제공하므로 레코드 규모가 증가할 경우 활용가치가 높다고 볼 수 있다. 연산자 및 필드 검색과 동시에 분류 제한 검색이 가능하도록 서



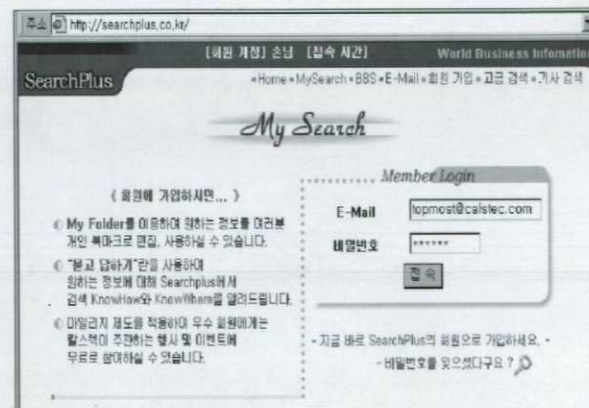
〈그림 6〉 키워드 검색 결과 화면



〈그림 7〉 고급 검색 화면

비스한다. 연산자는 AND(논리곱), OR(논리합), AND NOT(논리차) 연산을 지원하고 있다.

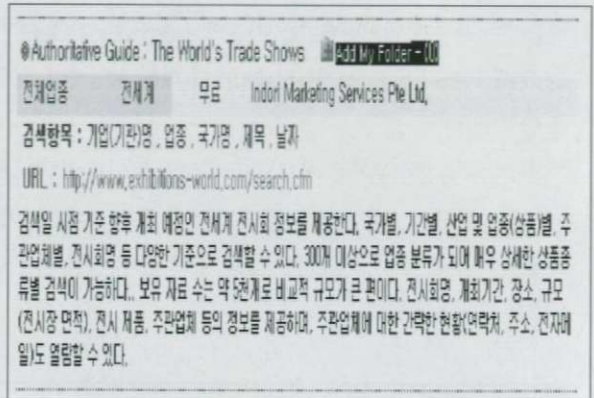
이 외에도 개인화 서비스를 통해 목록을 보면서 'Add My Folder'를 선택하여 개인 북마크 관리가 가능하도록 서비스하고 있다. 개인화 서비스로 이동한 모습은 〈그림 8〉과 같다. 이러한 개인화서비스는 검색자가 자주 이용하는 데이터베이스만 등록함으로써 다음 방문시 쉽게 자신의 '북마크 데이터베이스'를 찾을 수 있도록 제공하는 서비스이다.



〈그림 8〉 개인화 서비스 접속 화면

〈표 3〉 비즈니스 데이터베이스 분류

기업 및 금융정보	전화번호부/연락처검색
	기업 일반현황 검색
	기업 재무정보 검색
	주가 정보 검색
	환율 및 금융정보 검색
	기업 웹 사이트 디렉토리
동향/뉴스정보	복수 간행물 검색
	단일 간행물/저널 검색
	간행물 서지정보 검색
	기사초록 데이터베이스
	통신사 DB
법률 및 특허, 정부정책 정보	판례 데이터베이스
	법률 데이터베이스
	특허 전문(fulltext) DB
무역 및 지역 정보 검색	특허 초록(abstract) DB
	국가 정책 정보 검색
	수출입 업체 검색
	수출입 게시판 DB
	전시회 및 컨퍼런스 DB
인물 및 단체 정보 검색	국가 및 도시 정보 DB
	업종별 협회 및 NGO 검색
	전문가 DB
	인명록 검색
	전자메일 디렉토리
	인명 전화번호부 DB
상품 및 전자상거래	전자상거래 상품정보 DB
	온라인 경매 및 베품시장 검색
레퍼런스 및 다운로드	사전류 데이터베이스
	검색보조도구 - 주요 코드 검색
	도서 및 논문 검색
	웨어웨어 프로그램 데이터베이스
	음악/게임/이미지/영상 데이터베이스
	통계자료 및 조사보고서 검색
문화 정보	날씨 및 지리정보 데이터베이스
	스포츠 정보 데이터베이스
	공연/연극/문화 정보
생활 및 오락 정보	여행 정보 검색 서비스
	부동산 정보 데이터베이스
일반 검색엔진	채용 및 구인구직 데이터베이스
	웹 검색엔진
	웹 디렉토리
	유즈넷 검색엔진
	메타 검색엔진



〈그림 9〉 개인화 서비스 북마크 화면



〈그림 10〉 개인화 서비스 북마크 폴더 지정 화면

이러한 전문 데이터베이스를 활용할 경우 기업정보 뿐만 아니라 특허, 동향, 정부정책, 법률 등 다양한 분야의 비즈니스 정보 검색에 좋은 해결책을 제공받을 수 있다. 전문DB 디렉토리의 보유 DB 수는 1999년 9월 현재 3천개로 외국의 유사 서비스로 볼 수 있는 thebighub.com(2천개 DB수록)이나 search.com(1천개 DB수록) 보다 많은 '비즈니스 데이터베이스'를 수록하고 있다. 검색 가능한 데이터베이스 분류는 〈표 3〉과 같다.