

인터넷 정보검색기술 Patent Map 분석

Patent Map이란 특허기술정보를 분류·분석·가공·정리하여 필요한 기술정보를 쉽게 파악할 수 있도록 한 특허기술정보 해석서로서 특허청과 한국발명진흥회에서는 2000년부터 매년 24개 테마를 발굴, 최근 2003년까지 산업분야별 96개의 테마를 개발하여 CD-ROM과 Web Service를 통해 지속적으로 보급하여 왔습니다. 여기에서는 2003년에 기 수행된 PM보고서 중 전기/전자분야 「인터넷 정보검색」과 제에 대하여 기술의 주요 구성도, 기술 및 산업시장동향, 주요 특허동향분석 결과, 향후전망 등에 대하여 간략하게 소개하고자 합니다.

이 과제에 대하여 자세한 내용을 알고자 할 경우에는 Patent Map 홈페이지[<http://www.patentmap.or.kr/>]에서 무료 서비스를 시행하고 있으므로 많은 열람 바랍니다.

※ 문의사항은 아래 연락처로 하시기 바랍니다.

- Tel) 02-3459-2865-71

- Fax)02-3459-2879

기술의 주요 구성도

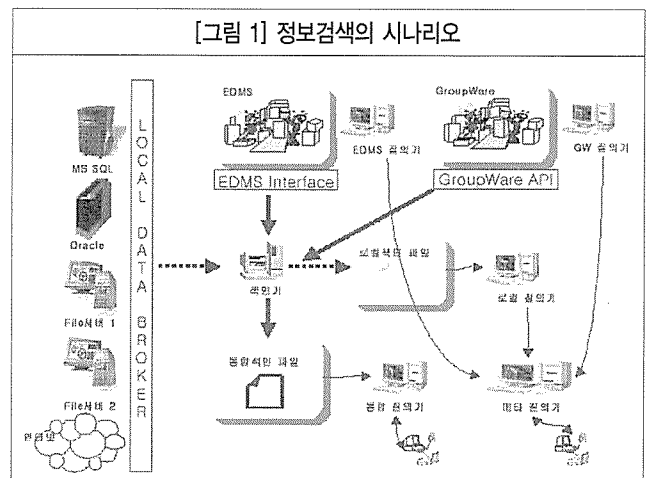
인터넷 정보검색 기술은 기존의 문헌 정보 검색 기술을 기반으로 하여 인터넷의 발달로 인해 생겨난 엄청난 양의 정보를 정확하게 검색하기 위해서 최근에 대두되기 시작한 기술이다. 이 기술은 사용자의 여러 가지 입력 방법을 통한 질의를 통해 질의를 분석하고 그 질의를 분석 및 확장하여 관련 문서정보를 추출하여 결과를 되돌려 주는 정보 검색의 일반적 과정에 네트워크를 이용하는 인터넷이라는 특성을 고려하여 세계 여러 곳에 분산되어 있는 정보를 효율적으로 검색할 수 있는 방법까지 포함하고 있다.

인터넷 정보 검색 기술을 위한 핵심기술로는 정확한 정보를 검색하기 위한 여러 가지 검색 기법, 검색된 정보를 효과적으로 통합하거나 분배하는 분류 방법, 검색된 정보를 사용자에게 시각적으로나 결과물로서 보기 좋고 활용 가능하도록 해주는 사용자 인터페이스 방법 및 정보의 효율적인 저장을 위한 저장 기법 등이 있다.

여기에서는 검색이론 및 기법, 정보표현 및 분류, 이용자 인터페이스, 저장기법으로 분류하였으며, 각각에 대해

세부기술로 또 다시 분류하였다. 분석대상 데이터는 1982년~2001년 사이에 출원된 해당특허 중에서 한국, 일본, 유럽은 출원공개 특허데이터를 기준으로 분석하였으며, 미국은 등록공개된 특허데이터를 기준으로 분석하였으며, 한국의 경우 KIPRIS를 미국, 일본의 경우에는 Delphion의 DB를 활용하였다.

[그림 1] 정보검색의 시나리오



산업 및 시장동향

정보검색 분야는 초기 목표였던 텍스트 색인과 유용한 문헌 탐색을 넘어서 비약적인 발전을 했다. 그리고 1990년대 초, 인터넷의 등장으로 전례 없는 규모로 생각과 정보의 공유를 가능하게 하였다. 이런 성공은 이용자의 컴퓨터 환경에 상관없이 항상 동일한 표준 사용자 인터페이스 착안에 기반하고 있으며, 결과적으로 사용자들은 통신 프로토콜이나 기계 위치, 운영체제로부터 자유로울 수 있었다. 이런 환경은 세계전역의 수많은 사람들에게 대단한 주목을 받게 되었고 많은 사람들이 컴퓨터를 사용하고 일과를 수행하는 방식에 있어 혁명을 일으키고 있는데 예를 들어, 홈쇼핑과 홈뱅킹이 매우 일반화되어서, 매출 면에서도 수억 달러가 발생되고 있다.

정보검색의 형성과 발전을 가능하게 한 사회적 환경은 크게 정량적 정보문제의와 정성적 정보문제의 둘로 나눌 수 있다. 다시 말하면 정보홍수라는 정량적 정보문제의 해결과 정보수요 및 정보요구의 분화라는 정성적 정보문제의 극복이 20세기 후반의 중요한 사회적 문제로 부각된 것이다. 이는 60년대 이후 지식 또는 정보가 국가의 생산력을 결정하는 요소로 작용해 온 사실에서 잘 나타나고 있으며, 정보의 상품화와 더불어 정보시장 또는 온라인시장의 형성이라는 산업구조의 변화가 이를 잘 증명해주고 있다.

이러한 사회적 환경이 미친 영향에 의해 형성된 정보검색의 새로운 경제적 환경은 오늘날 정보시장의 서비스와 생산물들의 다양화에 의해 특징지워지며, 앞으로도

새로운 유형들이 생겨나면서 다양화는 가속화될 것이다. 이는 온라인시장에 있어서 원하는 지식을 찾을 수 있도록 지원하고 도움을 주는 기능이 점점 더 중요하게 되다는 것을 의미한다. 특히 게이트웨이에 의해 생성된 부가 가치 기능은 최종이용자를 지향하는 온라인 시장의 확장에 많은 기여를 할 것이다.

국내 정보검색시스템 시장은 이제 시작단계라는 것이 전문가들의 중론이다. 구축 사이트도 일부 공공기관, 신문사, 방송국, ISP, 전자도서관 등 대량의 데이터베이스가 많이 구축되어 있는 기관들에 한정되어 있다. 한편, 국내 정보검색시스템의 구축 형태는 단순히 데이터베이스만을 구축하는 초기단계에 머물러 있다. 구축된 정보검색시스템의 검색방식 역시 아직은 단순한 키워드 검색이 주류를 이루고 있으며, 웹 상에서 무료로 서비스하는 경우가 대부분이다. 이처럼 국내 정보검색시스템 시장이 초기 단계에 머무르고 있는 주된 요인은 정보검색시스템에 대한 마인드 부족과 함께 수천만원에서 수억원에 이르는 구축비용이다.

최근 업계에서는 이러한 점을 감안하여, 저가의 정보검색시스템 보급에 주력하고 있다. 인터넷 사용자들이 급증하고 있으며, 이러한 사용자들이 손쉽게 접할 수 있도록 정보검색시스템의 광범위한 보급을 유도, 시장을 활성화시킨다는 전략이다. 업계 관계자들은 정보검색시스템 시장이 아직은 초기단계이지만 인터넷/인트라넷의 확산과 저가 검색엔진들이 잇달아 출시되고 있어 1~2년 이내에 시장이 활성화될 것으로 전망하고 있다.

[표 1] 검색연산 관련제품 현황

구분 제품명 비교내용	국내						국외		
	infoACE	다센21	레이다	소프트붓	마이다스	심마니	SEARCH97	FulCurm	BRS
개발회사	쓰리소프트	다센	한국정보공학	소프트와이즈	현인정보	한글과컴퓨터	Verity	FulCurm	BRS
공급사	쓰리소프트	다센	한국정보공학	소프트와이즈	현인정보	데이콤	쓰리소프트	삼보정보외	한미데이터
Boolean 검색	YES	YES	YES	YES	YES	YES	YES	YES	YES
Near 검색	YES	YES	YES	YES	YES	NO	YES	YES	YES
시소러스 검색	YES	YES	YES	NO	YES	NO	YES	YES	YES
필드별 검색	YES	YES	YES	YES	YES	NO	YES	YES	YES
날자조건 검색	YES	YES	YES	YES	NO	NO	YES	-	-
비교연산자 검색	-	YES	YES	-	NO	NO	YES	-	-
Wild Card 검색	YES	YES	YES	YES	YES	NO	YES	YES	YES
자연어 검색	YES	YES	YES	YES	YES	NO	YES	YES	YES
결과 재검색	YES	YES	NO	NO	YES	NO	YES	YES	YES
색인어 열람에 의한 검색	YES	YES	YES	YES	YES	NO	YES	YES	YES
구검색	YES	NO	NO	NO	NO	NO	YES	YES	YES
Multi-DB 검색	YES	YES	YES	YES	YES	YES	YES	YES	YES
DBMS 지원방식	DBgateway	YES	ODBC	YES	YES	NO	DBgateway	YES	ODBC

* 출처 : <http://www.dpc.or.kr/dbworld/document/9809/focus.html>

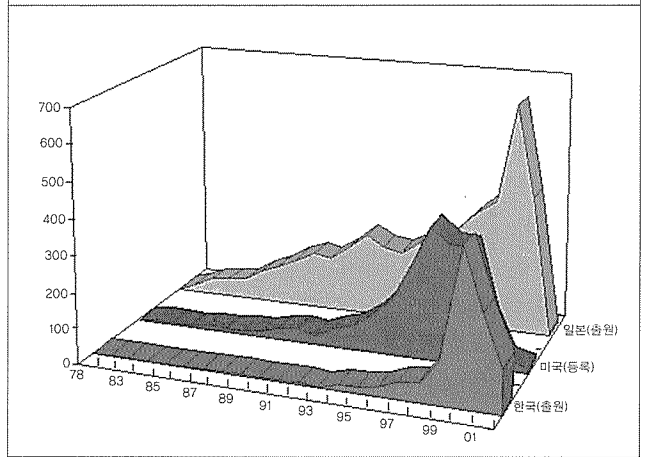
특허출원현황

인터넷 정보검색 기술은 크게 검색이론 및 기법, 정보표현 및 분류, 이용자 인터페이스, 저장기법으로 나눌 수 있다. 각 기술분야별로 세부기술을 살펴보면 검색이론 및 기법은 Search method, 다중검색, Multi-media 검색, 정밀검색, 의미검색을 포함하고, 정보표현 및 분류는 색인기법과 정보 범주화를 포함하며, 이용자 인터페이스는 정보 제시와 사용자 질의처리를 포함하고 저장기법은 Storage structure와 검색 시스템 구조를 포함한다.

인터넷 정보검색 기술의 특허 분석은 한국과 일본은 출원연도별 특허출원 동향에 대한 것이고 미국은 출원연도별 특허등록 동향에 관한 것이다.

한국의 특허출원동향은 1998년도부터 급격히 증가하는 경향을 보이는데 이것은 한국 내 인터넷 보급의 확산과

[그림 2] 국가별 연도별 특허출원(미국: 등록) 현황



국가적 사업인 국민 PC의 보급, 그리고 이전까지 모델을 이용하던 인터넷 사용자들의 전용선 사용 확대의 영향을

[그림 3] 기술발전도 요약 MAP

종분류	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00	01
Serch method																				
다중검색																				
Multi-media 검색																				
정밀검사																				
의미검색																				
색인기법																				
정보 범주화																				
정보제시																				
사용자 질의처리																				
Storage Structure																				
검색시스템 구조																				

받은 것으로 해석된다. 미국의 경우 1990년대 초 미국방성의 ARPANET 프로젝트를 시작으로 하여 인터넷의 발달에 따라 등록건수의 꾸준한 증가를 보이다 한국과 마찬가지로 1990년 후반에 이르러 전 세계적인 인터넷 보급의 확산과 비약적인 발전으로 현저히 눈에 띄는 등록건수를 나타내고 있다. 일본 또한 한국이나 미국과 마찬가지로 1990년대 후반에서 눈에 띄는 출원건수를 나타내고 있는 것을 알 수 있다.

기술분야별로 볼 때 세 국가 모두 인터넷 정보검색 분야의 기본이라 할 수 있는 검색이론 및 기법 분야에서 높은 출원건수를 나타내고 있으며, 특히 검색이론 및 기법 분야에서도 Multi-media 검색 기법과 의미 검색 기법은 향후 발전 가능성이 높아 출원(미국: 등록)건수가 증가할 것으로 예상된다.

인터넷 정보검색 기술의 기술발전도를 살펴보면, 다중검색 분야는 1980년대부터 꾸준히 발전하고 있으며 Multi-media 검색 분야와 정밀 검색 등의 기술은 90년대 이후부터 지속적으로 발전하고 있음을 알 수 있다.

향후전망

인터넷 정보 검색 기술의 검색이론 및 기법 분야에서는 차세대 웹으로 등장하는 시맨틱 웹이 활성화되면서 인터넷 정보검색 분야도 자연언어처리, 의미처리, 자동추출 및 번역 등에 관련된 특허들이 많이 등장할 것이다. 특히, 시맨틱 웹 구현을 위한 상위 계층 및 분야별 온톨로지가 등장할 것으로 예측된다.

정보의 표현 및 분류에 관한 분야에서 정보의 색인과 범주화는 점차 지능화 및 자동화될 것으로 예측된다.

이용자 인터페이스 역시 이용자 중심적, 이용자 친화적, 이용자 편리성으로 변화할 것이다. 정보공간의 확대에 따라 검색되는 정보의 양도 증가하기 때문에 어떻게 하면 이용자에게 효과적이면서 동시에 효율적으로 정보를 표현할 것인가에 관한 연구들이 증가할 것이다.

마지막으로 저장기법과 관련되어 축적되는 자료의 축적 구조는 여러 기법들이 개발되어 있지만 가장 기본이 되는 inverted file과 관련된 특허들이 지속적으로 등장할 것으로 예측된다. 특히, 검색시스템의 전반적인 구조와 관련된 통합적인 특허 즉, 인터넷 정보검색과 관련된 하드웨어와 소프트웨어를 통합한 인터넷 정보검색을 위한 플랫폼 같은 특허들이 출원될 것이다. ●

[그림 6] 인터넷 정보검색 기술의 발전 방향

