

# USPTO의 데이터베이스를 이용한 서치작업

김봉섭

특허청 사무관

## 1. 序言

미국특허청(USPTO)에서는 미국에서 등록된 특허에 관한 자료를 데이터베이스화하여 일반에 인터넷상으로 공개하고 있다.

특허청에서 심사관으로 근무하고 있는 필자는 평소, 특허출원된 출원서에 대한 심사업무에 상기한 데이터베이스를 적극 활용하여 많은 도움을 얻고 있다. 하지만 이 데이터베이스는 심사관뿐만 아니라, 특허분야에 종사하는 관계자, 특허출원을 하고자 하는 일반인·업체, 신기술 동향을 파악하고자 하는 연구원 등 다른 많은 분들을 위해서도 유용한 도구가 되기 때문에, 서치(검색)에 대한 체계적이고 상세한 설명을 통해 여러분들에게 도움을 주고자 한다. 참고로 필자는 아래의 주소로 이 글과 관계된 인터넷사이트뿐만 아니라 인터넷상에서 제공되는 특허와 관련된 여러 주소를 담고 있는 Bookmark 파일을 제공하고 있다. 독자여러분이 원한다면 온라인으로 바로 Bookmark 파일을 보내줄 것을 약속한다. PC통신 천리안 id:bongeun, 인터넷 e-mail주소: bongun@bora.dacom.co.kr

## 2. 홈페이지 찾아가기

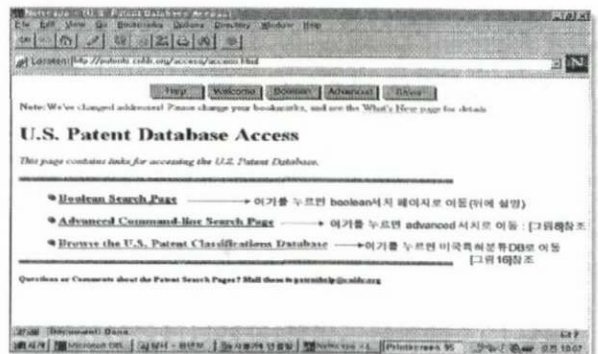
미국특허청에서 제공하는 초록검색 데이터베

이스는 상기한 Bookmark 파일에 등록되어 있는 홈페이지를 선택하거나(서치(search)▶미국특허청 제공-서치서비스▶홈페이지), 다음의 주소로 직접찾아가면 된다.

<http://patents.cnidr.org/acess/access.html>

접속이 되면 아래와 같은 화면이 뜨게 되는데, 여기에는 boolean서치와 advanced서치, 미국특허분류 데이터베이스의 세가지의 경우가 있음을 알 수 있다. 각각에 대해서는 아래에서 설명하기로 한다

〈그림 1〉 미국특허청 초록검색 데이터베이스 홈페이지



## 3. Boolean Search

원래 초록검색데이터베이스에는 advanced서치 외에 간단히 검색할 수 있는 툴로서, simple서치와 지금 설명할 boolean서치가 있었는데, 하나

의 필드에 대해서만 검색할 수 있는 simple서치의 경우 사용자의 활용도가 낮고, 다른 서치로도 충분히 검색가능하므로 데이터베이스 관리상 없어서 버렸고, 처음 쓰는 사람도 쉽게 사용할 수 있는 boolean 서치만 남게 되었다.

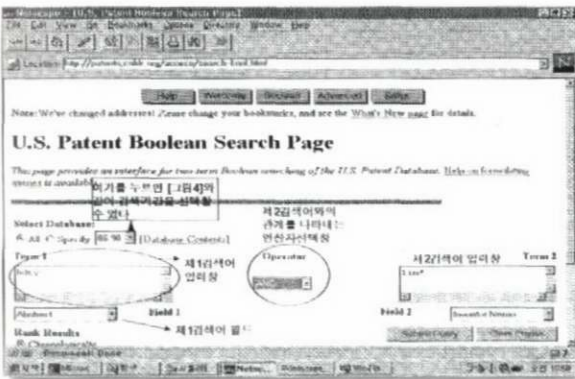
**boolean 서치의 장점**

검색어가 2개이하일 때 쉽게 검색할 수 있다. Advanced search의 검색구문을 외울 필요없이 검색하고자 하는 검색어와 검색필드만 알고 있으면 된다. 초보자라도 쉽게 사용할 수 있다.

**boolean 서치의 단점**

검색어가 3개 이상이 되면 검색이 불가능하다.

〈그림 2〉 초보자를 위한 boolean 서치 페이지(검색예)



〈그림2〉에서 검색필드 1로 초록(abstract) 필드와 검색어 "HDTV"를, 검색필드 2로 발명자 이름(IN)과 검색어 "lim\*"을 and시켜 놓았다. 이것의 의미는 발명자의 이름이 lim(한국발음 "임")으로 시작되는 모든 출원 중에서 초록(abstract)의 내용에 HDTV라는 단어가 들어가는 모든 문서를 찾으라는 것이다.(이런 검색명령의 배경은 얼마전 재미교포인 임재수 교수(MIT)가 다양한 포맷의 HDTV영상신호를 한 시스템에서 처리할 수 있는 특허를 미국에서 받았다

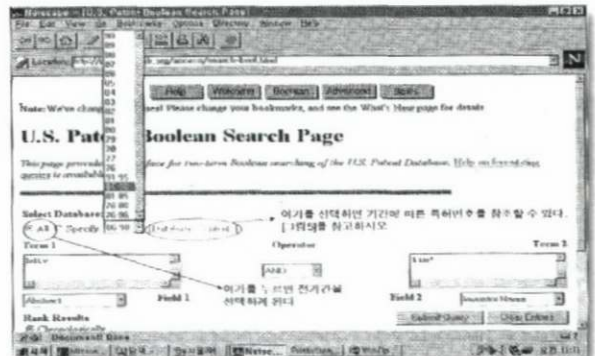
국내보도를 접하고 이를 확인하기 위한 것이다.)

〈그림 3〉 검색결과(뒤에 설명할 advanced서치의 결과와면 〈그림9〉과 비교하면 동일한 작업을 할 수 있음)



〈그림2〉에서와 같이 기간선택창을 누르면 다음과 같이 자신이 원하는 기간을 선택할 수 있다. 물론 기간 중 전체기간선택도 있다. 후자는 이 시점에서, 그러면 귀찮게 기간선택을 할 필요가 있는냐는 의문을 제기할 수도 있다. 그러나 우리는 인터넷을 하며 시간이라는 문제를 결코 등한시할 수 없기 때문에 최단시간으로 원하는 결과를 얻기 위해서는 가능하다면 기간도 선택하는 것이 바람직하다 하겠다.

〈그림 4〉 다양한 기간을 선택할 수 있다.(advanced search와 동일)



〈그림4〉에서 [Database Contents]를 선택하게 되면 아래의 〈그림5〉와 같이 기간과 특허번호



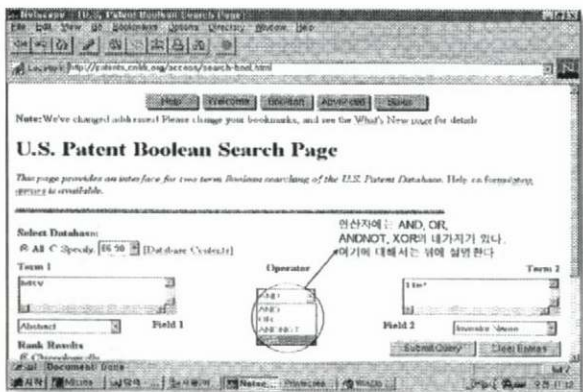
에 대한 관계를 참조할 수 있다. 여기서 알아두어야 할 것은 Utility는 일반 특허를, Design은 의장을, Plant는 식물특허를, Reissue는 재발행특허를 나타낸다. 이런 구분은 후에 설명할 advanced서치에서 특허번호(PN)으로 검색할 경우에도 동일하게 적용되므로 주의를 요한다.

〈그림 5〉 기간과 특허번호와의 관계(advanced search와 동일) - 참조 : 그림4

Database	Utility	Design	Plant	Reissue	Defensive Publication	NHE
95-96	5,317,359	D,351,359	P,9,006	RE,14,814		11,389
	1,569,040	D,341,110	P,9,646	RE,35,338		11,597
94	5,264,946	D,442,818	P,9,327	RE,34,493		11,370
	5,117,358	D,453,941	P,9,025	RE,35,413		11,368
93	5,175,886	D,342,170	P,9,081	RE,44,154		11,124
	5,248,885	D,342,817	P,9,026	RE,34,492		11,269
92	5,077,836	D,322,878	P,9,761	RE,33,786		11,897
	5,175,885	D,322,169	P,9,081	RE,34,123		11,123
91	4,980,927	D,413,201	P,9,408	RE,31,510		11,062
	5,077,835	D,322,877	P,9,760	RE,33,785		11,123

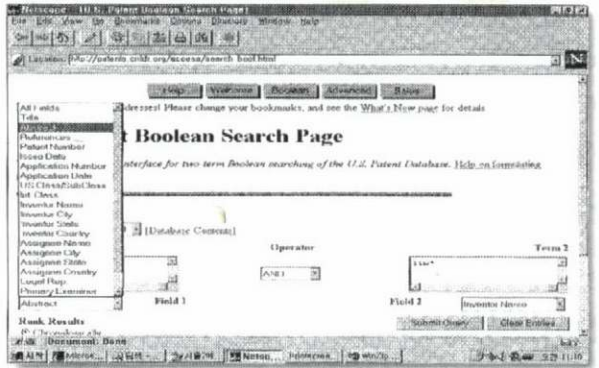
연산자는 검색필드 1과 검색필드 2 사이의 관계를 규정한다. 주로 AND와 OR를 사용하지만 ANDNOT과 XOR의 기능도 적절히 활용한다면 많은 도움을 받을 수 있다. 비록 boolean 서치의 경우엔 그 힘을 느낄 수 없지만 검색조건을 변화 무쌍하게 적용할 수 있는 advanced 서치에서는 그 막강한 파워에 놀랄 것이다.

〈그림 6〉 4가지 종류의 연산자 선택가능(advanced search에서 설명)



〈그림7〉에서와 같이 검색어의 검색필드는 "All field" (전체필드)를 포함하여 특허명세서에 나타나는 거의 모든 필드에 대해 개별적으로 선택가능하다. 영문으로 표기되어 있는 필드의 의미에 대해서는 advanced 서치에서 설명하고 있는 내용을 참고하라.

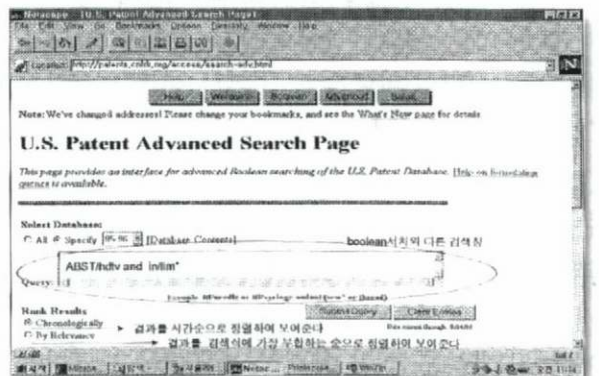
〈그림 7〉 출원서에 나타난 거의 모든 필드가 검색대상이 된다. (advanced 서치에서 설명)



#### 4. Advanced Search

Advanced 서치는 데이터베이스 검색에 익숙해진 사용자에게 아주 만족할 만한 결과를 가져다 주는 유용한 도구이다. 필자는 처음에 약간 번거롭고 힘들더라도 Advanced서치를 이용할 것을 권하고 싶다. 왜냐하면 참증을 찾기 위해 검색을 하는 경우 3개 이상의 키워드를 요구하는 경

〈그림8〉 Advanced search 초기화면



우가 대부분이기 때문이다. 차근차근 사용법에 대해 알아보기로 하자.

**장점** : 검색자 자신이 원하는 거의 모든 종류의 검색이 가능하다.

**단점** : 초보자가 사용하기 위해서는 약간의 암기사항 및 유의사항이 있어 숙련된 사용을 위해서는 시간과 노력이 필요하다.

<그림1>의 화면에서 advanced 서치를 선택하면 위의 <그림8>과 같은 화면이 뜨게 된다. 화면에서 보는 바와 같이 검색창은 하나로 구성되어 있어 원하는 검색식을 형식에만 잘 맞춘다면 아주 까다롭고, 찾기 힘든 검색도 원활히 해결할 수 있다. 기타 선택사항들은 boolean 서치와 동일하다. 다만, 앞에서 설명하지 않은 “Rank Results” 부분은 검색결과와 정렬순을 선택하는 것인데, “Chronologically” 를 선택하면 결과가 시간순으로 정렬되는 것이고(가장 최근의 것이 앞으로 옴), “By Relevance” 를 선택하면 검색 조건에 가장 잘 맞는 결과부터 정렬하여 보여주는 것이다.

검색식을 다 입력하고 <그림8>에서 “Submit Query” 를 선택하면 아래의 <그림9>와 같은 결과 화면이 나온다.

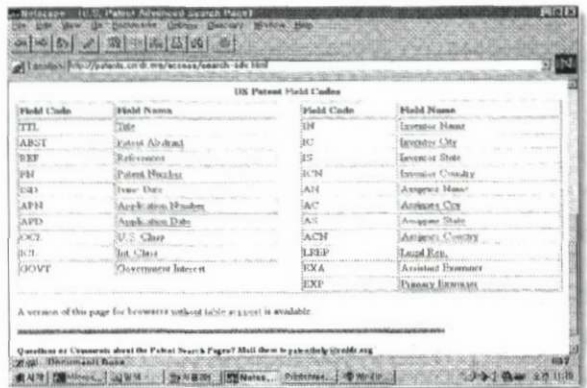
<그림 9> 검색결과(앞서 설명한 boolean search와 동일하다. 우리는 여기서 boolean search의 경우 advanced search와 같은 구문으로 변환하는 과정을 한 번 더 거침을 알 수 있다.)



### 필드를 이해하자

advanced 서치에서 필드를 이해하는 것은 효율적이고 올바른 검색을 위해 반드시 명확히 이해하여야 한다. 여기서는 필드에 대한 전반적인 모든 내용을 다 알아 보기로 한다.

<그림 10> 검색가능한 필드를 사용자가 쓰기 쉽게 하도록 하기 위하여 각 필드에 대한 설명과 사용예를 설명하고 있는 페이지 [http://patents.cndir.org/access/search\\_adv.html](http://patents.cndir.org/access/search_adv.html)



Advanced 서치로 검색가능한 필드는 도면을 제외한 전부라고 해도 과언이 아니다. 다시 말하면 검색식 사용에만 익숙해진다면 아주 치밀한 검색이 가능하다는 것이다. 다음은 검색가능한 필드목록으로 영문약어는 검색시 필요하므로 외워두는 것이 좋다(TTL(TiTLe ; 제목)

- ABST(ABBTract ; 초록)
- REF(REFErence ; 인용참증)
- PN(Patent Number ; 특허번호)
- ISD(ISsue Date ; 특허발행일)
- APN(APPlication Number ; 출원번호)
- APD(APPlication Date ; 출원일)
- OC(U.S. CLass ; 미국분류)
- ICL(Int. CLass ; 국제분류)
- IN(Inventor Name ; 발명자)
- IC(Inventor City ; 발명자 도시)
- IS(Inventor State ; 발명자 주 - 미국의 출원의 경우, cf. ICN)



ICN(Inventor CouNtry ; 발명자 국가 -  
미국출원의 경우 IS)

AN(Assignee Name ; 양수인)

AC(Assignee City), AS(Assinee State),

ACN(Assignee Country)

LREP(Legal Reprisentative ; 법정대리인)

EXA(Assistant Eaminer ; 보조심사관)

EXP(Primary Examiner ; (주)심사관)

검색식 사용은 처음에는 익숙하지 않아 번거롭고 힘든 작업처럼 보이나, 아래에 설명되어 있는 검색식 기재의 원칙, 연산자(operator) 등의 사항만 제대로 이해하면서 꾸준히 사용하다 보면 아주 쉽다는 것을 알 수 있을 것이다.

**검색식 기재의 원칙**

대원칙 : 필드명/검색식

예) ABST/HDTV, TTL/TUNER,  
IN/KIM\* 등

날짜의 경우(ISD, APD) : 다음의 몇가지 형태가 가능하다.

예) 적용날짜 96년 11월 6일

1. 19961106
2. 11-6-(19)96, Nov-6-(19)96,  
November-6-(19)96

\* (19)는 생략가능, 해당년월의 모든 날을  
원할 경우: \* 적용가능

예) 1995년 4월의 모든 출원에 대해 알고 싶은  
경우 : ISD/4--95

이름의 경우(IN, AN, LREP,EXA,EXP)

이름기재형식 : last\_name:first\_name initial

예) 발명자 중 김봉섭 을 찾고 싶은 경우 :

IN/ Kim; Bong s

문자의 경우(대소문자 구분없음)

1. 한 단어 : 따옴표(" ") 불필요  
예) HDTV 가 들어간 제목 : TTL/hdtv
2. 두 단어 이상 : 따옴표를 반드시 써야 함  
예) HDTV튜너가 들어간 초록 :

ABST/ "HDTV tuner"

**연산자의 종류**

AND

예) HDTV AND TUNER : HDTV와  
TUNER가 모두 들어간 것

OR

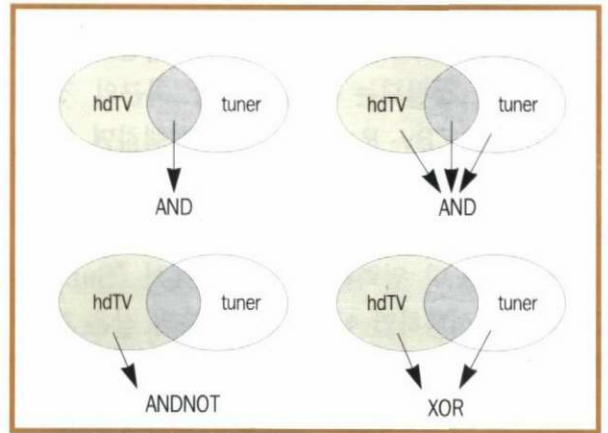
예) HDTV OR TUNER : HDTV와  
TUNER 중 하나라도 들어간 것

ANDNOT

예) HDTV ANDNOT TUNER :  
HDTV가 들어간 것 중에 TUNER제외

XOR

예) HDTV XOR TUNER : HDTV나  
TUNER 중 하나만 들어간 것



**필드서치**

PN

예) 특허번호 5444498를 찾고 싶은 경우 :  
PN/5444498

\* 의장(design)의 경우 번호 앞에 D를  
붙인다. : PN/D2496514

식물특허의 경우 번호앞에 PP(plant  
patent)를 붙인다. : PN/PP9046

ICL

예) H04N7/01의 경우 :

1. ICL/H04N-7/01
2. ICL/H04N 7/01

**기타**

Wild card(right truncation : \*)

예) IN/KIM\* : 이름에 KIM이 들어가는 발명자 전체

**조합**

예) ICN/KR AND ABST/television and ABST/tuner and ABST/auto\*

**주의사항**

“ABST/television and ABST/tuner” 와 “ABST/television and tuner”의 차이는 무엇일까?

차이를 알기 전에 먼저 알아 두어야 할 것은 바른 검색식의 기재는 ABST/television and ABST/tuner와 같이 동일한 필드의 검색일 경우에도 반드시 검색어(여기서는 television, tuner) 앞에 검색필드(ABST)를 적어야 한다는 것이다.

위의 두 검색식 중에서 두 번째 식은 사용자가 잘못 기재를 하였거나, 숙달된 사람이 의도적으로 적은 두 가지의 경우를 상정해 볼 수 있다. 여기서 의도적이란 말은 위의 두 번째 식이 다른 의미를 가진다는 것을 뜻한다. 이 경우에서 처럼 tuner앞에 특정의 필드를 기재하지 않으면 모든 필드에 대해서 검색한다는 것을 의미한다. 하지만 대부분의 경우, 잘못 이해하고 위와 같이 사용하는 경우가 많고 이런 경우 누차 강조하지만 모든 필드를 검색하므로 시간손실이 크다는 것을 알아야 하겠다.

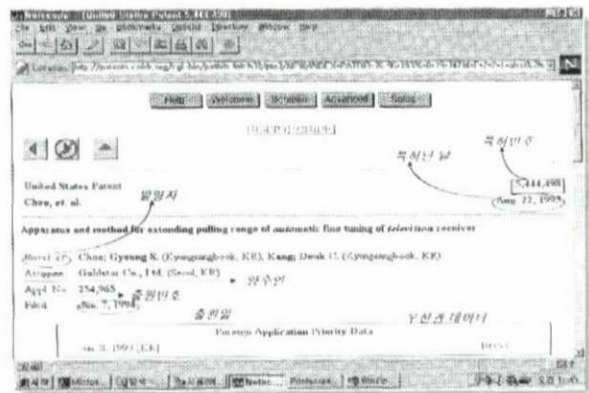
참고로 이 글에서 사용하고 있는 검색식 ABST/television and tuner and auto\* 대신 ABST/television and ABST/tuner and ABST/auto\*를 사용한다면 결과가 달라짐을 확인할 수 있을 것이다.(2건의결과가 1

건으로 될 것임)

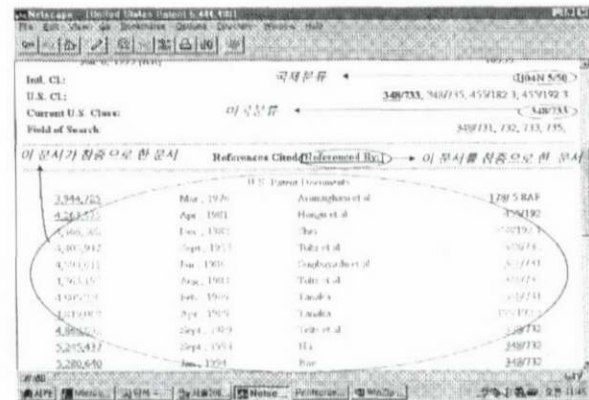
**검색하기**

어떤 임의의 검색식을 이용하여 얻어진 결과는 아래의 그림에 표시된 것과 같은 많은 정보를 포함하고 있다. 그림을 보면서 하나하나 체크해 보라!

〈그림 11〉 검색결과1



〈그림 12〉 검색결과2

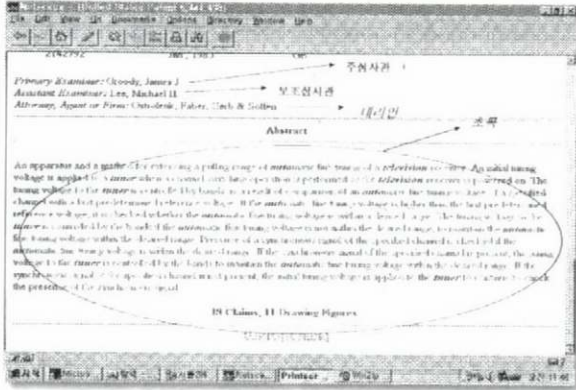


〈그림12〉 검색결과 2를 에서 미국분류를 클릭하면 해당 미국분류로 넘어가게 된다. 또 〈Referenced By:〉라고 써진 부분을 누르면 이 문서를 인용한 다른 출원서의 리스트가 나타나는데, 이는 우리가 지금 사용하고 있는 PFS검색과 비슷한 개념이다. 그리고 제일 밑에 리스트형식



으로 나타나는 내용이 이 출원서의 상세한 설명에서 인용한 문서들에 대한 정보로, 해당 번호를 누르면 그 문서로 넘어가게 된다.

<그림 13> 검색결과3



위의 그림에서 보는 바와 같이 마지막 부분은 담당심사관의 이름과 대리인 그리고 초록으로 구성되는데, 초록의 내용에는 청구항과 도면의 개수에 대한 정보도 제공하고 있다.

**검색식에 따른 검색속도 검토**

아주 많은 자료를 검색하다 보면, 동일한 결과지만 그 결과를 얻어내는 시간은 검색식을 얼마나 효율적으로 쓰는가에 따라 상당한 차이를 보일 수 있다. 한 가지 예로 동일한 결과를 얻는 다음의 두 검색식에 대해 검토해 보기로 하자.

먼저 첫 번째 검색식에서는 1차로 389건을 얻고(ABST/tuner and television), 이 자료와 다음 조건(auto\*)에서 두 번째 결과인 69건을 얻은 후 이 결과를 최종적으로 ICN/kr와 조합하여 2건의 검색결과를 구해낸다.

하지만 두 번째 검색식에서는 동일한 과정에 의해 1차로 28건을, 2차로 7건을, 최종적으로 2건의 결과를 얻고 있다.

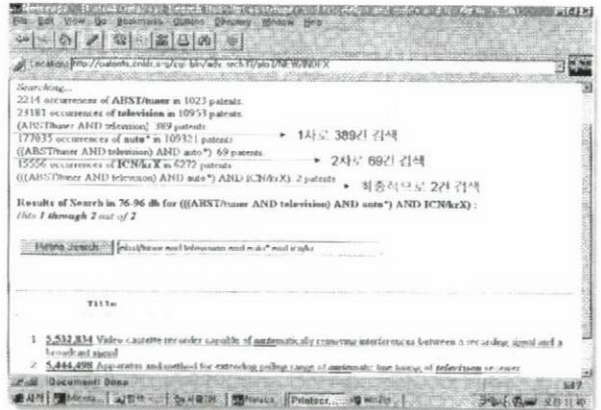
당연히 두 번째 검색식이 빠르다는 것을 알 수 있다.

그렇다면, 이러한 차이는 어디에서 기인하는 것일까?

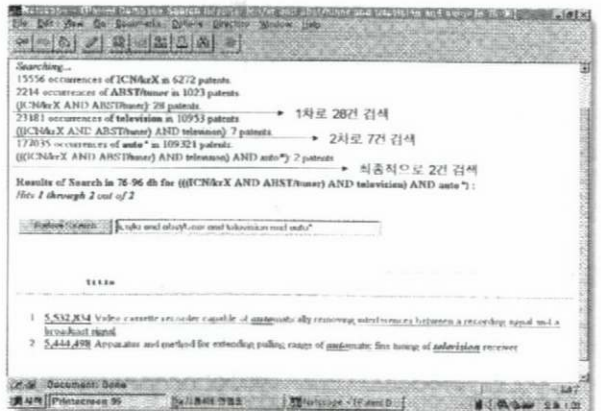
이유는 간단하다. 컴퓨터는 우리가 입력한 검색식을 왼쪽에서 오른쪽으로 훑어 나가기 때문이다.(물론 괄호가 없을 경우에만 해당한다) 따라서 우리는 1차검색에서 검색결과를 최소화하기 위해 가장 적합한 키워드(상황에 따라 달라질 수 있기 때문에 원칙은 없다)를 앞에(왼쪽에) 위치시키는 것이 필요하다.

본 예에서는 ABST/tuner and television (389건) 보다 ICN/kr and ABST/tuner(28건)가 더 효율적인 방법인 것으로 판단된다.

<그림 14> 검색식에 따른 검색속도차이1(검색은 왼쪽에서 오른쪽 키워드 순으로 된다)



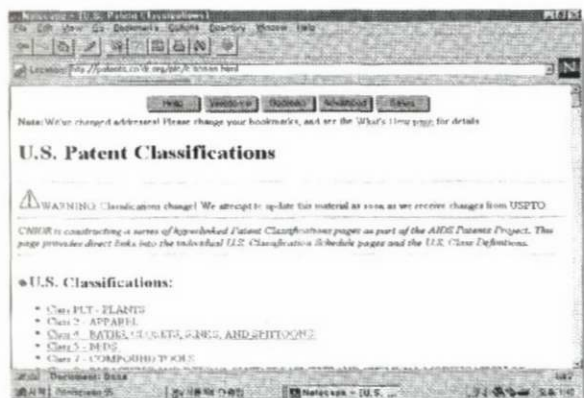
<그림 15> 검색식에 따른 검색속도차이2(앞의 검색보다 진보된 검색; 키워드조합의 중요성을 말함)



## 5. 미국특허분류 데이터베이스

홈페이지 <그림1>의 세가지 항목 가운데 제일 마지막에 있는 미국특허분류 데이터베이스에서는 미국특허청이 자체적으로 운영하고 있는 미국식 분류를 통한 자료검색이 가능하다. IPC체계를 따르는 우리로서는 큰 효용을 바랄 수는 없지만, 미국식 분류를 검토하므로써 모호한 기술내용을 명확히 파악할 수도 있고, 검색된 문헌에 대한 불명료한 부분을 보완할 수 있는 효과가 있어 나름의 의미를 부여할 수 있겠다. 아래 그림은 초기화면으로, 상기한 <그림12>에서 미국분류를 누르면 해당하는 미국분류로 넘어가게 된다.

<그림 16> 미국분류 데이터베이스의 초기화면



시 조속한 시일내에 전산화를 완료하여 국민에 양질의 서비스를 할 수 있으리라 기대해 본다.

그런 점에서 미국특허청에서 전세계를 상대로 이렇게 안정된 검색도구를 제공하고 있다는 점은 우리에게 시사하는 바 무척 크다. 우리도 시간이 지나면 저런 멋진 시스템을 구축할 수 있으리라는 희망찬 생각을 하면서 비록 우리의 것은 아니지만 잘 활용하였으면 하는 마음 간절하다. **DrC**



## 6. 結 語

우리 특허청은 계속해서 전산화 작업을 한 결과, 미국특허의 도면까지 검색할 수 있는 전산망을 일부 구축 완료하였다. 참으로 고무적인 일이 아닐 수 없다. 하지만, 서치를 해 본 심사관이라면 아직도 완벽하고 안정한 검색시스템이 되기 위해서는 많은 시간이 소요될 것이라는 생각을 했을 것이다. 또한 미국과 같은 검색서비스를 제공하기 위한 데이터베이스 구축은 아직도 요원하다는 판단을 하였을 지도 모른다. 하지만 우리 역