advantage of this valuable information resource.[5,6]

LITERATURE CITED

(1) K. M. Donovan and B. B. Wilhide, "A Users Experience with Searching the IFI Comprehensive Data Base to U.S. Chemical Patents", *J. Chem. Info. Comput. Sci.*, following paper in this issue.
(2) S. M. Kaback, "A User's Experience with the Derwent Patent Files", *J. Chem. Info. Comput. Sci.*, **17**, 143 (1977).
(3) R. J. Rowlett, Jr., private communication
(4) "CAS Today", pamphlet, Chemical Abstracts Service, Columbus, Ohio, 1974.
(5) G. O. Platau, "Documentation of the Chemical Patent Literature", *J. Chem. Doc.*, **7**, 250 (1967).
(6) D. B. Baker, F. A. Tate, and R. J. Rowlett, Jr., "Changing Patterns in the International Communication of Chemical Research and Technology", *J. Chem. Doc.*, **11**, 90 (1971).

A User's Experience with Searching the IFI Comprehensive Database to U.S. Chemical Patents†

K. M. DONOVAN* and B. B. WILHIDE

Pennwalt Corporation, King of Prussia, Pennsylvania 19406

Received April 20, 1977

The use of the IFI Comprehensive Database to U.S. Chemical Patents to provide patent information services to an R&D staff is described. The history of the database, its present structure, and their effects on search techniques is presented. A new method of searching tapes of the vocabulary and USPO classification text by a string search program provides the search staff with quick access to these search aids. Experiences in training new searchers to use this database are given. Costs of operating the file are presented.

The IFI Comprehensive Database to U.S. Chemical Patents (IFI database) is the result of a merger of the IFI Uniterm Index and the Du Pont Index to U.S. chemical patents. These two files were merged in 1971.

The Uniterm Index was developed in 1955 as a coordinate term index. It was first published in printed form as a dual dictionary which allowed the searcher to perform simple Boolean logic intersections by comparing the inverted index lists of the dictionary for matches. The vocabulary was open ended and uncontrolled. In addition to the dual dictionary of major terms, a single dictionary of minor terms was provided for index entries with low postings. The minor terms were primarily chemicals. In 1961 the Uniterm Index was published in magnetic tape and sold with a software package that used a weighted term retrieval technique. The searchable elements included the assignee and the major terms of the dual dictionary. The minor term section of the index was not issued in computer format. In 1964 the USPO classifications for the patents were added to the file. Subscriber demand also produced a controlled vocabulary and a review of the minor term vocabulary for new candidates to the major vocabulary.

The Du Pont Index to U.S. Chemical Patents was begun in 1964. A controlled open-ended vocabulary was used for both chemical and nonchemical concepts. A chemical fragmentation scheme was designed to index both specific chemicals and generic chemical structures. A system of roles was developed to specify whether the chemical indexed was present, reacting, or a reaction product. An elaborate role scheme was also designed for the indexing of polymers.[3]

The Comprehensive Database was created in 1971 when the IFI/Plenum Data Co. acquired the Du Pont Index and merged it with their Uniterm Index. This database now contains 400 000 U.S. chemical patents dating back to 1950. The file is updated quarterly. Lag time from patent issuance to appearance in the database is six months. The growth rate is approximately 20 000 patents a year.

Table I. Combinations of Search Elements Used

Search elements	No. of searches
Terms and/or Compounds (T C)	198
T C/Class Code/Fragments	77
T C/Class Code	60
T C/Fragments	43
T C/Assignee	24
T C/Assignee/Class Code	14
Assignee	12
Class Code/Fragments	8
Fragments	5
Class Code	5
Assignee/Fragments	3
Assignee/Class Code	1
T C/Assignee/Fragments	1

The selection of patents for the database is done on an automatic basis for a specified list of USPO subclasses. The Official Gazette is scanned for patents of chemical interest appearing in other subclasses.

The searchable elements of the database are: accession number, patent number, assignee, USPO subclass OR and XR, compounds, chemical fragments, and general terms. The accession number indicates the year of issue. Both the patent number and accession number can be used to retrieve the complete index record for a patent. The index elements can be used in any combination in search strategies. Table I presents the variety of combinations of search elements used in designing our search strategies.

The assignee is that appearing on the face of the patent. This requires the searcher to investigate the name, history, acquisition, and merger history of any company to be searched. At the present time a major revision of the assignee vocabulary is in progress. Name changes such as Pennsylvania Salt Manufacturing Co. to Pennsalt Chemicals to Pennwalt Corporation are being collected into a standard format. However, patents issued to companies which have merged or been acquired are not posted to the parent company. An example would be the patents issued to Sharples, Wallace &

† Presented in the symposium on "Meeting the Challenges of the Changing Patent Literature", Division of Chemical Information, 173rd National Meeting of the American Chemical Society, New Orleans, La., March 21, 1977, and the 11th Middle Atlantic Regional Meeting of the American Chemical Society, Newark, Del., April 20, 1977.

*Author to whom correspondence should be addressed.

Tiernan or other companies acquired by or merged with Pennwalt. These appear under their own names in the database. There are 23 000 assignees on the database.

The USPO classifications assigned to the patent are searchable. Both original reference (OR) and cross reference (XR) subclasses and main class are searchable. These classifications represent subject indexing of the patent by the Patent Office. They are used as concepts in the construction of search strategy. Another valuable use of the patent class and subclass is to partition the file into subsets. This saves CPU time in searching the database. The classifications are updated annually to reflect the classification revisions in the USPO.

The General Term Vocabulary is controlled and open ended. It includes general chemical concept terms, trade names, nonchemical terms, and search term only (STO) terms which reflect old indexing philosophy. This vocabulary contains 10 000 terms. The nonchemical concept terms make up approximately 25% of this vocabulary. The chemical terminology in this vocabulary includes generic class terms, trade names, general polymer terms (this is the designation IFI uses for terms like poly(vinyl chloride)), and the discontinued indexing terms called "Search Term Only". The latter are terms from the old IFI Uniterm vocabulary which were very general terms, such as "Organic Chlorine Compounds" or class terms, such as "Styrene, Styrenes". The "Search Term Only" category must be used in designing strategy to search the database for the 1950–1964 time period. It is also profitable to search these terms up to the merger date of 1971.

The Compound List Vocabulary is a register of chemicals which have five or more patents indexed to them. There are 14 000 compounds in this section of the vocabulary. The compounds on this list are also indexed with the chemical fragment scheme which permits substructure search of this list. You may set up a substructure search of these 14 000 compounds so that the answers are automatically incorporated into your search strategy. Thus, if you want your search to include all compounds in the vocabulary that are amino alcohols, you do not have to manually search the vocabulary list. A fragment search can be designed that will extract all such compounds from the 14 000 and insert them into your search. You do have the option of first searching the 14 000 compounds and printing out the list selected for review. The 14 000 compounds can also be searched when novelty questions are asked concerning a class of compounds.

The chemical fragment indexing scheme[4] is applied to three types of chemical structures: the Markush formula; indexer generated generics; and specific chemicals to be indexed which are not on the Compound List. The chemicals listed in the minor term index from the 1950–1964 Uniterm Index have been indexed with the fragmentation scheme. This work was completed with the 1975 update. Four types of fragment terms are used: (1) atoms present, (2) functional groups, (3) configuration descriptors, and (4) ring descriptors. A "Must-Possible" approach is used. All fragments applicable to every member of the generic family described are indexed with both "Must" and "Possible" forms of the descriptors. The fragments which describe the alternative components of the structure are indexed with the "Possible" forms of the descriptors. At search the "Possible" fragments are asked for and the "Must" fragments which do not apply to the structure being searched for are negated. The "Must" fragment list contains 151 structure and configuration codes that are available for this negation technique. The judicious use of negation enables the searcher to eliminate irrelevant answers. We also at times use the "Must" fragments for searching. This is useful in searching for compounds composed of highly posted chemical fragments. The searcher knows that there will not be complete recall on this type of inquiry. In addition to "Must-Possible" indexing, a link technique is used to prevent scrambling of the fragmentation indexing to two or more chemical structures from the same patent.

Roles are used with index terms which describe chemicals to indicate whether the chemical is present, a reactant, or a reaction product. A special role, "00", is applied to the Uniterm indexing. An extensive role system defines polymers. The roles reveal the class of polymer formed and whether it is a homopolymer or copolymer and also whether the polymer is modified. The roles define seven classes of polymers: polycarbons, polyesters, polyethers, polycarbonates, polyurethanes, polyamides, and polysiloxanes. An eighth role group is used to index other polymers. The roles are applied to the prescribed monomer chemical terms. Prescribed starting monomers are indexed for polymers to allow the search to recall all references to particular polymers without regard for variety in starting materials. If a polymer process is claimed, the actual starting monomer is indexed also. Polymers are also indexed with generic class terms and terms from the general term vocabulary. If the monomer is not on the Compound Term List, it is indexed by the fragmentation scheme and the appropriate polymer roles assigned. The indexing of class 1, the polycarbons, will illustrate the use of the polymer roles. The number "one" appears in all the roles used with this class. The roles used for unmodified polycarbon homopolymers are 41 (present), 51 (reactant), and 61 (product). If the polymer is modified the roles are the inverse −14, 15, 16. For copolymer polycarbons the role group is 71, 81, 91, 17, 18, 19. An example of the indexing for the production of unmodified poly(vinyl chloride) follows:

Vinyl chloride	role 61
Vinyl homopolymers/halide	role 30
poly(vinyl chloride)	role 30

For a modified copolymer of vinyl chloride and vinylidene chloride the indexing will be:

Vinyl chloride	role 19
Vinylidene chloride	role 19
Vinyl copolymers/halide	role 31
Vinyl chloride≅vinylidene chloride copolymer	role 30

The search aids provided by IFI include the General Term Vocabulary, Thesaurus, Compound Term List, Fragment Term List, and Assignee List; frequency tables for the postings in each year; Indexer's and Searcher's Guide; and a manual describing the computer programs and input format.

Other aids we use are the USPO Manual of Classification and its index, the U.S. Patent Classification—Subclass Listing, and the Patent Number Sequence Classification Record.

At the 1976 users conference, 3M reported on the use of a character string search program against the vocabulary tapes. We have also adopted this technique for search of the vocabularies and found it very useful in preparing search strategies. Another use we have made of our string search program is to run it against the text of the USPO classification which IFI supplies with its edit program. This has been very helpful in identifying subclasses for search. It has proved an important educational tool in training searchers to use the USPO classification. The string search program permits the use of "AND and "OR" logic. A special feature of the program used in searching the patent classification allows the searcher to specify a hierarchical relationship between the strings to be searched.

This database permits the recall of indexing for the searcher's review. When we have prior art available, we have the complete indexing for it printed out. This is a great help in building search strategies. Since indexing vocabulary has

changed significantly during the history of the database, it is very helpful to look at the indexing of patents in the older section of the file. This is also an important teaching tool in the education of new searchers.

The computer program supplied to operate the database is in four parts. Part one is an edit program which validates the search input against the vocabulary tapes. The second part is the search program which operates against the compound list fragmentation indexing and the full database. The third program sorts the output according to instructions given in the input. The fourth program is the report writer. The edit, search, and report writer programs are frequently operated independently as well as in conjunction with a full search retrieval request.

The retrieval system is based on weighted term searching.[1,2] In weighted term searching the index terms are grouped to logically express the various parts of the search inquiry. The terms in each group are given individual positive values. Each group is assigned a threshold value (a minimum acceptable weight, MAW). The sum of the values of the weighted index terms must meet or exceed the MAW in order for the patent to be retrieved. There is an implied "AND" between the groups. The weighted term search technique has proved to be most valuable. It allows the searcher to control the relative importance of search terms to the strategy. In the output the total weight for each patent retrieved is given. This enables the searcher to evaluate the hit in relation to the search strategy. A unique feature in the IFI system is an answer frequency table. This is available as an option for each search. Frequency tables for USPO subclasses, assignees, and total answer weights can be requested. The USPO classes and subclasses show the concentration of answers in the subclasses. An evaluation of this frequency table may indicate unknown subclasses which should be incorporated into search strategy. The assignee frequency table will highlight the companies most active in the field of the search. An evaluation of the total weight frequency will reveal the strengths and weaknesses of the search strategy. If an unprofitable search term causes a high percentage of irrelevant hits in the answer list, the total weight frequency table enables the searcher to identify it.

Boolean logic AND, OR, AND NOT can be simulated using the weighted term technique. As mentioned earlier, AND logic exists between the groups of terms. An OR relation exists between each term within a group that has a weight value equal to or greater than the MAW. The AND logic can be forced between two sets of terms within a group by assigning values less than the MAW to these terms but which together add up to a value which is equal to or more than the MAW. A list of synonyms in a search strategy can all be identified by the same weight value by assigning the value of zero to the second and subsequent synonymous terms in the strategy. The individual terms in the synonym list will not be identifiable in the search output. You will only know that one or more terms in the synonym list were present in the indexing.

The AND relationship between groups can be altered by use of the search chaining technique. Within a group of related requests a chain can be established. There are two types of chained outputs. In the first, all searches identified on the chain eliminate duplicate answers from the searches higher on the chain. This is used to eliminate the review of duplicate answers found in related search strategies. In the second chain technique, all answers found in the chain of searches are printed in the output of the first search in the chain. With this technique it is possible to establish an OR relationship between groups.

If the MAW for a group is set at zero and the term weights are assigned a value of −1 the group becomes an AND NOT logic parameter in the search strategy.

When the MAW is set at zero and the term weights are positive, the group does not enter into the logic of the search. However, any answer hits which contain indexing given in the zero group will show these term weights in the total weight which accumulates for the hit. This technique is used frequently when patents of a company are reviewed. Known technology is signaled in a zero group so that review time can be concentrated on the unknown aspects of the company's technology. Another use for this technique is to review the content of a specific subclass. It also permits the searcher to see additional indexing to the retrieved patent which is not essential to the logic of the search but which may help evaluate the answer hit.

Our average answer list is 200–300 hits. However, we do review output with as high as 3000 hits by means of the weighted term technique. A user survey conducted in 1976 also found that the number of answers per search averaged 245.

An average of three strategies is prepared for each search topic. A separate strategy is usually prepared for the old IFI Uniterm indexing. Both broad and narrow strategies are designed for the searches. The chaining technique is used for these multiple strategies so that redundancy in the output is removed. Searching the file in-house permits unlimited use of alternate strategies. New searchers are especially encouraged to prepare alternate strategies. The search results are screened by the searcher. We use the total answer weight and title as the first screen. The CA abstracts are reviewed. We have on file all U.S. chemical patents since 1965 on microfilm. Copies of the most pertinent patents are supplied to the client. The computer printout is given to the client for his evaluation and use. We maintain copies of the search strategies for two years.

The search output can be sorted by accession number, patent number, assignees, patent subclass, or total weight. Alternate sorts of the same answer list can be done. Sorts within sorts are also available, i.e., major sort by assignee with each assignee list in descending patent number order.

The output can also be produced in tape form. This is called a "save tape" which can then be used as a database tape for searches. We construct save tapes in areas of technology that we search repeatedly. The CPU time saved by searching these mini databases is considerable. We are now working on a small information system for one of our research projects using the IFI database. A save tape of all pertinent U.S. patents is being prepared. A file of patent copies is being collected. The save tape will be used as a search database for this file.

The use of the IFI database in the research environment must be placed in the context of our total information work. The Information Services Department serves the research, development, product application, engineering groups, and technical management at the Pennwalt Technological Center. This site is the corporate research center for the company's chemical interests. We also provide limited information services for the Patent Department and technical managers at headquarters and plant locations.

The services offered to our Technological Center staff include current awareness programs, literature and patent searches, and general library reference. We have a staff of eight people, five professional and three clerical. At present two professionals are trained to operate the IFI database. A third member will be trained this year. It has been our experience that it takes one year to train a chemist to use the database effectively. This does vary depending on the person's experience in working with computer databases. Training is done one on one. During the first three months, strategies for each request are discussed with the trainee. The trainee is

Table II. Comparison of Search Resources

Resources used	No. of requests
IFI and on-line and manual	31
IFI and on-line	44
IFI and manual	12
IFI	71
On-line and manual	35
On-line	129
Manual	63

encouraged to do alternate strategies for search requests and to learn to use the search aids. As proficiency grows the trainee is encouraged to work independently and usually by the end of six months can handle routine requests. The chief difficulty in training new people to be effective searchers has been in educating them to appreciate the history of the database and to teach them that effective searchers must consider all the indexing options: the old uniterm indexing, the USPO classification, as well as the current Du Pont indexing.

We have found that attendance at the annual user's conference provides an excellent opportunity for users to exchange information on search techniques and applications of the database.

Our retrospective search service for literature and patents is performed on demand. We are responsible for searching and gathering together the appropriate corpus of literature for any questions we work on. Evaluation and analysis of the literature is done by the individual researcher.

In 1976 we processed 385 requests for searches: 302 were done for Technological Center staff, 54 for headquarters staff, and 29 for other locations. Of the 385 requests, 158 required the search of the IFI database. A comparison of the use of the IFI database with other sources is given in Table II.

The 158 IFI searches required 451 requests against the IFI database. Rounded, this gives an average of three inquiries for every search request. At present one person works about one-third time on IFI searching.

Our primary use of the IFI database is in performing prior art searches for the research groups. Requests for searches come to us at the idea stage, during the conduct of research, and at the time an invention record is prepared. Ideally, searches should be made at each of these steps, and we encourage our people to do this. Our Patent Department expects us to do a novelty search of the IFI database before an invention record is submitted. Our project and group leaders also have on-going state-of-the-art searches updated semiannually and annually on topics related to their research programs. Another use for this database is to review competitors' patent activity. Recently we have been experimenting with the use of the database for our marketing staff. An example of their use is the review of patent activity related to our chemical intermediate type products. The use of the reaction role on the chemical provides access to information of use to them.

The computer we use for this database is an IBM 370/158 with the VS1 operating system. Our access to the computer is via a remote job entry terminal (IBM 3776 Terminal). We have an ideal arrangement for operating the search work on this database. There is no restriction on the time of day or day of the week the batch can be run. Input is keypunched by our own clerical staff. The edit program and string search program are operated on a disk pack as separate modules. Our turn-around on the operation of these programs is less than five minutes. After the batch is edited it is entered for running against the search, sort, and report writer programs. The output is printed on a line printer at our location. When we operate the program during the day our turnaround time is about two hours. We usually submit the search work in the late afternoon and take it off the printer in the morning. We operate very small batches of three to eight searches per batch. The CPU time for batches of this size averages six minutes. The IFI programs are designed to operate at a limit of 36 searches per batch. Most users do save their work until they have large batches to run. It is more economical for those operations which are charged back for specific computer runs. We are not charged back for our specific work. Computer services for our division are operated as a division overhead service. We do keep account of our costs. Our 1976 cost for operating the in-house file was $231 per search topic ($151 for subscription cost and $80 for computer cost). The cost per search strategy (we average three strategies per search topic) was $80. The IFI service bureau charges $250 per search strategy. In 1976 the IFI/Plenum subcommittee of the Manufacturing Chemists Association's Committee on Technical Information[5] designed a test batch of eight questions which was run by four comprehensive database subscribers to compare computer costs. The cost results on this test batch were reported by three users. The test was run in two installations on an IBM 370/165. At Gulf Research and Development Co., the CPU time was 7.2 min and the cost was $102.25. At Eastman Kodak the CPU time was 0.264 min and the cost was $125.38. We ran the test batch at Pennwalt on an IBM 370/45. Our cost was $260 for 44 CPU min. As noted earlier we are now using an IBM 370/158 for this database. Problems encountered using the IFI database are the following:

1. The many changes in vocabulary and indexing philisophy and technique are stumbling blocks to learning the effective use of the total database.

2. The fragmentation scheme produces a high level of false retrieval when used to search for structures composed of highly posted fragments such as halogen, acid, ester, ether, and hydroxy functional groups.

3. The earliest prior art in a search can be missed if a broad search strategy is not used. This occurs in any controlled vocabulary situation. There is always a lag between reality and indexing vocabulary.

The unique features of this database which we believe make it a valuable information resource for the R&D environment are the following:

1. Chemical Structure Searching. The chemical fragment system permits the searcher to query the file for a specific structural moiety without regard for a hierarchy of functional groups as in Chemical Abstracts manual searching. The substructure search of the compound list permits the use of many compounds in a search strategy without the burden of naming each of them in the strategy.

2. Search of the USPO Subclasses. This database puts the USPO classification system for chemistry in-house. Searches on topics for which no subclasses exist are possible in this file. In effect the searcher can create a "subclass" of patent art related to his topic. It is possible to search hundreds of subclasses at once. There is no need to restrict USPO subclass search to what is physically possible as in manual searching.

3. Polymer Chemistry. The unique system used to index polymer chemistry enables the searcher to specify the level at which he wishes to search: (a) search can be done at the monomer level, (b) search can be done at the chemical fragment level, (3) search can be done at named polymer level, (4) search can be done at the generic level.

4. Time Saved. This is a very subjective opinion. We have had experience with this database in a "one man shop". Having the database in-house meant we could turn around substantial amounts of search work. The order of magnitude was about tenfold compared to manual searching. The database is the first used in any patent search. The results obtained in the quick turnaround from the IFI data are used

to decide whether other searches are needed. We can turn our IFI patent search requests around in 24 hours if that is needed. The database is used for all levels of searching: "quick and dirty", "find some background information", novelty search, validity search, infringement search, and state-of-the-art search.

We have been subscribers to this database since its creation in 1971. Prior to that time we were subscribers to the Uniterm tape and the printed Uniterm Index. We have had the advantage of growing with the file. Our decision to purchase the Comprehensive Database in 1972 was based on our need to have superior patent coverage in the fields of polymer chemistry and synthetic organic chemistry and also to permit more cost-effective use of professional time in providing search services. We also believed that the price of the merged database was an information bargain. Our experiences of the past five years have not altered our original position. The present annual subscription cost is high, and we do monitor our costs of operation each year to evaluate the cost effectiveness of this database to our total information service. However, intangible benefits are realized in having any database in-house; i.e., the search work is held confidential, the searcher is very familiar with the information needs of the client, and the searches can be restructured or up-dated at a reasonable cost. These factors also are considered in maintaining this database in-house.

In conclusion, we believe the IFI Comprehensive Database is a cost effective means of providing our R&D staff with a fast, comprehensive in-house search service for the United States patent literature.

LITERATURE CITED

(1) J. M. Cattley, et al., "Retrieving Patents by Weighted-term Searching", *Chem. Eng. Prog.*, **62**, 91–96 (Oct 1966).
(2) P. T. O'Leary et al., "Computer Searching of Chemical Patents", *J. Chem. Doc.*, **5**, 233–238 (1965).
(3) L. E. Rasmussen and J. G. Van Oot, "Operation of DuPont's Central Patent Index", *J. Chem. Doc.*, **9**, 201–206 (1969).
(4) M. Z. Balent and J. M. Emberger, "A Unique Fragmentation System for Indexing Patent Literature", *J. Chem. Inf. Comput. Sci.*, **15**, 100–104 (1975).
(5) Private communication, John W. Lotz, IFI/Plenum Data Co.
(6) Private communication, Patrick T. O'Leary, Gulf Research and Development Co.

A User's Experience with the Derwent Patent Files†

STUART M. KABACK

Information Research and Analysis Unit, Research Corporate Services, Exxon Research and Engineering Company, Linden, New Jersey 07036

Received April 20, 1977

The Central Patents Index-World Patents Index System, produced by Derwent Publications, Ltd., provides a wide range of products and services, ranging from alerting and in-depth abstracting through retrospective information retrieval. All chemically related patents have been covered since 1970, and nonchemical patents since 1974, but portions of the file go back as far as 1963. Multiple tools using a variety of parameters are available for information retrieval, and the recent availability of on-line search capability has increased substantially the utility of the Derwent files. In summary, the CPI and WPI constitute a powerful information resource which works superbly in some areas, less effectively in others.

The family of Derwent's patent information services forms one of the most complex information systems offered to the public today. There are alerting services in the form of expanded titles, short abstracts, and long abstracts, multiply packaged and repackaged. There is retrieval via classified browsing files, various printed indexes, punch card sorting, computer tapes, and, since early 1976, on-line interactive searching. In the on-line files alone the subject retrieval parameters include Derwent classes, manual code classes, international patent classes, multipunch codes, and title keywords. Wherever you approach the Derwent system there are generally multiple options—options that may tend to confuse even the most experienced Derwent user, let alone the uninitiated.

Derwent's Central Patents Index and World Patents Index cover virtually all of the world's chemically related patent publications, as well as a large percentage of nonchemical patents. Table I shows something of the volume and character of this patent coverage. Derwent's estimates for the current year put the annual total, chemical and nonchemical, at about 500 000 patent documents. Nearly half of these are "basics"—patent cases new to the data base. Most of the remainder are equivalent to references already in the system, although a small proportion are patents from so-called "minor" countries which are given partial coverage. About half of all

Table I. Derwent Patent Coverage, 1977 (est)

	CPI	WPI only	Total
Total basics	104 000	127 500	231 500
Total equivalents	147 000	100 500	247 500
"Minor" country patents not fully in system			16 000
Grand total	251 000	228 000	495 000

Table II. CPI-WPI Country Coverage

Full Coverage					
Belgium	(BE)	West Germany	(DT)	Soviet Union	(SU)
Canada	(CA)	France	(FR)	United States	(US)
Switzerland	(CH)	United Kingdom	(GB)	South Africa	(ZA)
East Germany	(DL)	Netherlands	(NL)		
Full Chemical Coverage					
Japan	(JA)				
Partial Coverage					
Brazil	(BR)	Israel	(IL)	Rumania	(RU)
Czechoslovakia	(CS)	Norway	(NO)	Finland	(SF)
Denmark	(DK)	Austria	(OE)	Sweden	(SW)
Hungary	(HU)	Portugal	(PT)		

† Presented in the symposium on "Meeting the Challenges of the Changing Patent Literature", Division of Chemical Information, 173rd National Meeting of the American Chemical Society, New Orleans, La., March 21, 1977, and 11th Middle Atlantic Regional Meeting of the American Chemical Society, Newark, Del., April 20, 1977.