A Patent Searcher's Personal Chronicle: 40 Years in the Evolution of a Profession‡

Stuart M. Kaback†

Information Research & Analysis, Research Support Services, ExxonMobil Research & Engineering Co., Annandale, New Jersey 08801

Received August 22, 2001

Forty years have passed since this chemist elected to attempt to become an information chemist, thus pursuing a career path he had not heard of during his undergraduate and graduate education. Much has changed during that period. Punched cards, microfilm and microfiche, coordinated term indexes, and more have come and gone. The US Patent and Trademark Office has issued more than 3 million patents, exceeding its total output in all the years that came before. Online database searching replaced prior reliance on printed indexes and classified card files, and now seeks to redefine itself to stand up against the juggernaut of the Internet. In the face of all that change the traditional abstracting and indexing function is still with us, though not without considerable reshaping. This paper surveys this landscape of change and suggests that if we are wise, we will nurture this intellectual activity far into the future.

Slightly more than 40 years ago, in January 1959, as a young chemistry graduate student approaching the end of my Ph.D. research, I took a chance and decided to attempt a career in information research. I was not looking to escape the laboratory; I loved it. And I would not characterize myself as having been an exemplary user of the literature as a graduate student. But the work sounded intriguing, and I decided to give it a try.

It took longer than I expected to finish my graduate work, and I didn't actually start that career until April 18, 1960. Except for a brief side step into the lab I've remained in information research since then. During that time incredible change has taken place. I like to think that I've contributed in a positive way to some of that change.

Looking back at the profession I entered 40 years ago, it's remarkable to see how primitive it was in so many ways. I believe that it's instructive to look at some of the conditions that prevailed in the world of technical information back in 1960, and during the ensuing years.

The day after I started work the US Patent Office (its name didn't yet mention trademarks) issued patents starting with 2,932,826. This address was delivered on August 24, 1999, also a patent issuance day in the US, beginning with 5,940,-878. More than 3 million patents had been issued since April 1960, more than in all the preceding years.[1] The number of patents issued that April week in 1960 was 906, while in 1998 the average was just about 2900 per week, and in 1999 the weekly yield on occasion topped 4000. Back in the early 1960s it was quite common for US patents to take four, five, six, or more years to issue. They took so long to issue that I made no attempt in those days to follow them. Rather, I became intrigued with foreign patents, because they tended to appear much more quickly, providing early technological leads that I could share with my colleagues, whether in the lab, in management, or in my own group. Today, of course, many US patents issue in less than 2 years, and late in 2000 the US was finally planning to begin publishing many US applications in 18 months, in keeping with the rest of the world.[2]

For another prehistoric condition consider the indexing of *Chemical Abstracts* (*CA*). The only subject indexes to *CA* then were the annual subject indexes; it wasn't until 1962 that those indexes became semiannual. The problem was that those annual indexes were late; if memory serves correctly they got to be about two years late before the situation started to improve. The simple keyword indexes to *CA* didn't start to appear in individual *CA* issues until 1963. Finding recent information in those days was a challenge, to say the least.

How did we find anything recent? I can't speak for other organizations, but at what was then Esso Research we had an in-house abstracting group, supplemented by cottage labor, that prepared bulletins for distribution to company personnel. The abstracting program served multiple purposes. Our information researchers reviewed the abstracts before publication, to select highlight items that were featured on the front page of the bulletins. We spotted discrepancies that might appear in an abstract. When an abstract seemed to be in an area of particular interest to one of our colleagues, we sent a copy to that colleague. The abstracts were classified according to a detailed system, each one getting at least one subject classification. The bulletins were cut up, the abstracts were pasted on 4 by 6 cards, and the cards were filed by subject, by author, and by patent assignee. We searched by pulling out file drawers and leafing through the cards. When a category got to be too voluminous, it was usually reclassified, but we never did get around to reclassifying the lubricant additives category. Thank heavens I never searched in that hypervoluminous area in those days.

The abstracts were generally good ones, informative rather than indicative, following the teachings of the late Ben Weil, a past Skolnik winner I had the privilege to have as a colleague. Literature abstracts were normally made from full documents, although some were prepared from abstract sources. Some patent abstracts were made from complete

† Corresponding author phone: (908) 730-2559; e-mail: stuart.m.kaback@exxonmobil.com.

‡ Herman Skolnik Award Address, ACS 218th National Meeting, New Orleans, LA, 8/24/99.

10.1021/ci010088f CCC: $22.00
Published on Web 01/18/2002

specifications, but some were prepared from abstracts, which ranged from rather good ones provided by Derwent to some wretched ones in patent office gazettes. In general our abstracts tended to be short of the kind of detail we find so valuable in many of Derwent's better Documentation abstracts, or in the best abstracts from *Chemical Abstracts*. We pored through the *CA* annual indexes when they arrived, and searched them year by year, grateful when a collective appeared.

Economics doomed our in-house abstracting program during the 1960s, but replaced it with something exciting: the American Petroleum Institute's (API) indexing system, which was developed in Ben Weil's group by a team headed by Gerry Jahoda.[3] A year's sample of petroleum refining and petrochemical literature was sampled in order to develop the indexing vocabulary, a faceted hierarchical system in which an indexing term posted all broader terms, and a search for a term retrieved all references indexed by that term or by narrower terms. At first we searched it mainly by printed tools: a dual dictionary, similar to the *Uniterm Index*, to coordinate two indexing concepts; an alphabetic subject index showing the titles of documents indexed by a term, and—if that term was deemed a major descriptor for the document—showing in addition the remaining index terms. Soon we were doing batch computer searches; how well I remember learning how to operate an IBM card punch to prepare my search strategies. I think that few of us were unhappy when the API databases went online in the mid-1970s.

Turning back to *Chemical Abstracts*, patent coverage in *CA* during the 1960s was quite limited. For a very small handful of countries there was supposed to be complete coverage of patents that described new chemistry. I always found that criterion to be interesting; it seemed to me that it was up to patent examiners and not abstractors to determine what was new. In any event, given the subject selection criteria, all relevant patents were covered consistently only for the US, Germany, and the UK; theoretically for France, but there were gaps; and for one or two other countries sometimes, but not all the time. It didn't help at all when the roof caved in at the Belgian patent office during the days when Belgium was a full-coverage country, and one of the key sources of early patent publications dried up for quite a while.

In most other countries that were covered, such as Japan, only patents to nationals were abstracted. Most countries then only published patents that had been determined by examiners to be patentable—The Netherlands was not to teach the world the dubious beauties of universal publication of applications, with deferred examination, until 1964. There were certainly chemical inventions—who knows how many—that issued only in countries for which Chemical Abstracts Service (CAS) followed the "nationals only" policy, hard-core chemical patents that never issued in one of the full-coverage countries—perfectly legitimate chemical inventions that were never covered by *CA*. *CA*'s document analysis in those days was based on what was actually done in examples, with hard data; information found in the claims was ignored. It could have been worse: I have been told on what I believe to be good authority that there was in those days a body of opinion at a high level of CAS management that believed that patents were of no value, and shouldn't be covered at all. I can't absolutely confirm that, but it's a frightening thought. After all, the *Engineering Index* has never covered patents, despite the heavy involvement of various branches of engineering in invention and innovation. We should be grateful for what we do have from that period, with a special tip of the hat to Dr. Russell Rowlett, who, among other things, got CAS pointed in the right direction when it came to patents.

The Dutch began the shift to universal publication of patent applications in 1964. Germany did the same in 1968, with the added fillip of pouring out over the course of a few years their huge backlog that had caused them to switch to universal publication in the first place; at least the Dutch never saddled us with that. It took Japan, though, to really turn the patent world on its head by coupling universal publication with the Japanese practice of having very narrow patent claims scattered over multiple documents, rather than multiple claims in a single document. They started their new system in 1971, and soon they were publishing 100 000, 200 000, nearly 350 000 Kokai each year. All this along with their unique numbering system based on the year of whoever was Emperor at the time; the patent world has never been the same. (Note: Japan abandoned the Imperial year for granted patents in 1995, and has done so for unexamined applications (Kokai) in 2000, but still retains the Imperial year in the numbering of patent applications.)

It wasn't until nearly 1980 that we got the European and Patent Cooperation Treaty (PCT) applications and patents, which had among other things the virtue of replacing multiple equivalent patent documents with just one. A negative outcome of the increasing popularity of PCT filing is the fact that most PCT publications today designate more than 100 countries, creating a situation in which counts of country coverage have become essentially meaningless.

I'm getting ahead of myself; let's return to the early days. Patent documents in the 1960s were far less informative than they are today. In 1960 a patent claiming foreign priority would tell where and when, but would not cite the priority application number. That wouldn't start happening until 1962, with the French leading the way. International filing wasn't as prevalent then as it is today, but it was certainly significant, and files of patent equivalents began to spring up in various places. Our approach at Esso was to maintain a card file of abstracts from our abstract bulletins, based on priority date and country, and a bright young lady with a high school education attempted to match similar inventions claiming the same date. It worked remarkably well, so well that our file formed the foundation for the API's patent concordance. Some of it is still functioning today as part of the Derwent database, thanks to the merger of the API and Derwent files. As one might guess, there were times when a multiplicity of similar inventions on a given date overloaded our clerk's capabilities. I still remember fondly untangling 22 related cases on Ziegler catalysts filed by one company on the same day back in the late 1950s, lots of fun to straighten out if you're the kind of nut I am about such things.

Another useful bit of information, one that we take for granted in today's patents, is the listing of examiner's citations. The US had started to list those in the late 1940s, back in the days when an examiner would normally list a small handful of highly relevant items. Many years were to pass before the introduction of the applicant's Information

Disclosure Statements (IDS) would induce examiners to cite everything on an IDS, giving us today far too many patents with 50, 100, more than 200 citations—all of them presumably close to the new invention, but not close enough to invalidate it. These voluminous citation lists are in my opinion information pollution. Just as swollen lists of designated states on PCT applications make it impossible to draw meaningful conclusions from patent country coverage data, so do swollen citation lists strip much of the value from what could be extremely useful information. The move to endless citation lists is especially unfortunate in light of the fact that the USPTO has chosen not to categorize relevance levels, as do the European Patent Office and the various PCT searching authorities.[4] A dump of a complete file can provide the information needed to answer questions, but provides it in useless form; mountains of indiscriminate information have little or no value. Selectivity is essential. This is one of the unfinished tasks on my list of things to try to change before I close up shop.

Step by step the type of patent front page we're familiar with today has evolved, providing us with more and more useful information. I do wish, though, that something could be done about patent attorneys and agents who still saddle some patents with meaningless titles such as "Chemical Product", and supplement these with wonderful abstracts such as "An additive for a lubricating oil comprises a functionalized polymer." In my ideal world such applications would be declared a priori unpatentable.

The decade of the 1960s was quite a watershed for users of chemical information, and especially for users of patent information. I have already mentioned some of the substantial milestones of that decade: the issue keyword indexes in *Chemical Abstracts*; the API indexing system; the enhancement of priority data on patent front pages, initiated by France, and the trend to rapid universal publication of patent applications with deferred examination, begun by the Dutch and ultimately followed by virtually all industrialized countries, with the US finally joining in (in part) in 2000. But how much more did that decade bring? Eugene Garfield's *Science Citation Index*. The earliest keyword-in-context (KWIC) indexes, perhaps best known from *Chemical Titles*. The first programs for computer searching of the IFI *Uniterm Index*, developed at Gulf Oil. The start of the CAS *Registry*, by teams under the management of Dale Baker and Fred Tate. *CBAC* and *POST*, experimental machine-readable databases from CAS. Bill Wiswesser's line notation, which played a big part in the information ferment of the 1960s, becoming a key feature of systems ranging from ISI's *Index Chemicus* to many in-house databases. The Chemical Notation Association is gone now, but the Chemical Structure Association, centered in the UK, assumed part of its mantle and is still doing valuable work today. The sixties brought the first of the Derwent documentation services, *Farmdoc* and *Ringdoc*, *Agdoc*, and *Plasdoc*, the entrepreneurial ventures of Monty Hyams that formed the basis for today's *World Patents Index* database. And, as I mentioned earlier, the new focus of CAS on patent information, one of the accomplishments of the leadership of Russell Rowlett, who took over as editor in 1967.

The name-dropping has not been unintentional; it has included quite a roster of past winners of the Skolnik award, all of whom played key parts in the explosive progress of the sixties. It was my privilege to be able to call each of them a friend. I know that I have learned and benefited from each of them.

I'll mention some other prehistoric conditions that have come and gone during my shift. Photocopiers. We actually had a very early, nonfunctioning Haloid Xerox copier when I started out, but used mostly Kodak Verifax wet copiers. They made prints that curled up if not restrained, and discolored badly. Others used 3M's Thermofax thermal copiers, not wet but also prone to discoloration and curling. Once the more modern Xerox machine and its competitors arrived, along with highly detailed Derwent documentation abstracts, we suddenly found that we didn't have to write out all of the details of a report. We learned quickly how effective it could be to write a little explanatory text around copies of Derwent documentation abstracts. And despite today's ease of downloading abstracts by computer, this troglodyte for one still clings to some of the old methods, which in many instances seem to be able to provide more useful information, although the current availability of Derwent extension abstracts online certainly makes downloading more effective than it was. The extension abstracts are in any event only available for a limited period; there's no satisfactory replacement for documentation abstracts with older references.

Returning to the historical survey, information's stone age saw searches that used IBM punch cards, edge notched cards sorted with needles, and optical coincidence cards. Microfilm and microfiche and their relatives have come and gone, doing valuable service along the way. CD-ROMs replaced them, and are still in use but seem already to be headed out to pasture. Dual dictionary searches of the *Uniterm Index* and other files, which allowed the coordination of two indexing concepts, and were not at all friendly to the use of groups of alternative or synonymous search terms, seem quaint today, but they were once the best we had. We had sequential batch searches on magnetic tape. And then there was online, that anachronism that many seem ready to bury.

I can remember the early days of ORBIT, from System Development Corporation (SDC), when we could print five answers at a time, and had to keep telling the computer that we wanted to keep going. I can remember when Dialog didn't provide for nesting, and when as late as the mid-1980s they still had a 50-character limit to a query, limiting a fragment search in the *CLAIMS Comprehensive* database to no more than four terms. When ORBIT first allowed us to cope with the API linked term system by arranging linked terms in alphabetical order so that we could SENSEARCH them, a unique form of stringsearching within a sentence unit. I can remember when the results of a *CA Registry* dictionary search had to be rekeyed to search them in *CA* files, before SDC listened to us and developed PRINT SELECT, and Dialog followed with MAP. Now that we can search in multiple files, now that we can tell the system to go and search selected terms in other databases, a capability enhanced by different hosts but especially by STN, it's easy to forget how very far we've come. I can even remember when the subject-searchable information we had online for *CA* was limited to title and issue keywords; no deep indexing, much less searchable abstracts.

The whole concept of applying information developed in one database to a search in another database is one that has

always been close to my heart. It doesn't take very long for any searcher to recognize the value of using more than one database to address a problem. Had I not learned this lesson in any other way, it was made crystal clear by the results of a careful study carried out in the early 1980s by the API's Patent Task Force. The members of that Task Force carried out about a dozen searches in a series of databases: *Chemical Abstracts*, Derwent *WPI*, *CLAIMS Comprehensive*, and *APIPAT*. We collaborated on and polished search strategies so that each database was being used to its best advantage, to the best of our abilities. In brief we learned that not one of the databases produced a perfect search for any of the queries; no one found all of the known set of relevant hits. A paper I presented in the early 1980s, on the realities of online patent searching, reported on these results, and dreamed of a utopia in which we could combine the strengths of different databases to produce superior search results.[5] Ultimately the online hosts have made crossfile searching possible in many situations, and ultimately ORBIT and Derwent provided a real utopian combination, the merged *Derwent−APIPAT* database.[6] Crossfiling and the merged file have given us tantalizing pieces of that pie in the sky. Will we ever get more? Note, incidentally, the difference between crossfile searching, in which searchable parameters from different databases can interact synergistically, and multifile searching, where a search is run in more than one file, but there is no interaction between the files. Crossfile searching has the potential to be significantly superior to multifile searching.

It is worth noting that the merged database was conceived at a meeting of the API's Patent Task Force. As the former long-term chair of that Task Force I am quite unable to pinpoint whose idea it was, beyond being certain that it was not mine. But the origin of the species isn't my reason for mentioning it. One of the glories of the API system over the years was a series of Task Forces that guided the journal coverage, the vocabulary development, and the patent coverage of the system, and monitored closely the quality. Sadly, not much of that system of Task Forces remains today, a victim of cost-cutting. Close interactions among database producers or online hosts and knowledgeable users have in my experience have been invaluable in developing, maintaining, and improving high-quality value-added products. I have been privileged to participate in advisory groups for several information provider organizations, and have seen first hand the benefits. I have also seen how corner-cutting has chipped away at some of these forums. The result is never beneficial.

Which brings us to my recommendations, summarized in Figure 1. The first is that database producers and online hosts need to work closely with small and knowledgeable user groups. This is certainly not to say that they shouldn't talk to the masses. It doesn't in any way imply that anyone should be disenfranchised. But there is something special that happens when a manageable-sized group, preferably international, gets together and kicks ideas around. It's much more productive than kicking vendors or users around.

Another modest proposal. I referred in passing to STRINGSEARCHing and SENSEARCHing, an antique method of locating data elements that are in a particular sequence. Until online link operators were developed, SENSEARCHing formed an effective, if primitive, method of utilizing API's linking capabilities. This isn't the only system I've heard of

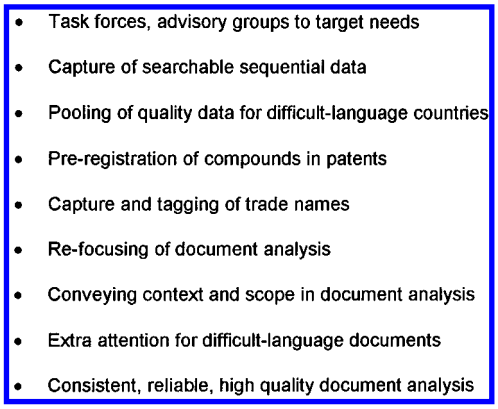

Figure 1. Suggestions for value-added information providers.

that stores and retrieves useful information by means of specific sequences; in the early 1970s one Herman Skolnik—now that is an interesting name to invoke at a Skolnik Award Symposium—Herman developed what he called the Multiterm system at Hercules, another information system in which sequences were important.[7] And if I'm not incorrect, sequences have a certain importance in dealing with genetic information, too.

For years I have talked of the need to be able to distinguish among sequences in such areas as catalyst manufacture, hydrocarbon conversion systems, and the like. Much of the focus of chemical information system development has been on ways to handle chemical structure, and rightly so, of course. But there has been very little focus on being able to distinguish a catalyst made by mixing A and B, then contacting them with C and finally with D, from one made using a different sequence of the same components. Yet the two may have very different properties, and may be the subject of different patents. There has been very little focus on distinguishing among chemical engineering inventions in which subprocesses are carried out in different sequences. Couldn't some system be developed in which a linked set of terms could be created in the specified sequence? One would narrow the search by searching for the individual search terms, then STRINGSEARCHing the sequence. How about it, somebody?

I mentioned the combined Derwent−API database, the product of cooperation between two organizations. Cooperation isn't always seen in today's competitive atmosphere. I can recall back in the 1970s when CAS and Derwent shared Japanese coverage for two years. The experiment died. Neither was satisfied with the output provided by the other. But the amount of information found in patents from countries with hard-to-handle languages continues to grow. Japan has been a major challenge, China and Korea are becoming ever more important, and the pool of competent document analysts who can provide meaningful document analysis in idiomatic and technically correct English does not seem to be inexhaustible. Is there not some way in which a pool of high-quality patent translations could be shared by Derwent and CAS, at least, with each one selecting and analyzing the key information according to its own precepts? There's a lot of subpar stuff out there today; cooperation could really help.

From time to time I've made the radical suggestion that it would be wonderful if the main patent offices worked out an arrangement with CAS in which the chemical substances

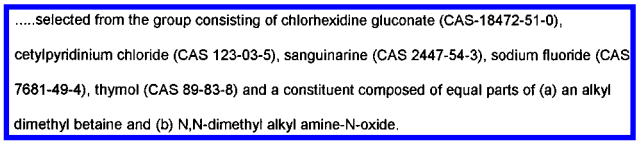

Figure 2. CAS Registry numbers in Claim 1 of US5869436.

in patent applications were processed by the *Registry* system, providing unambiguous access to the chemical substances in patents and anticipating the ultimate registration of new substances once the patents issued. In this light I was extremely interested recently to come across US 5,869,436, whose first claim includes five CAS Registry numbers, as shown in Figure 2. The full specification is full of CASRNs. I don't know the source, presumably the patentee, but I maintain that the idea has validity. Even better would be a system in which the CASRNs could be linked to role indicators identifying the functions of the substances in the patents.

Which brings me to my next suggestion. My colleagues and I use patents in a number of ways. We look at them to see if our inventions are novel. We look at them to see if someone else's patents might hinder our commercial operations. We look at them to see if patents of others or of our own organization are really valid. We look at them for research ideas. And we look at them for technical intelligence. One of the invaluable inputs to technical intelligence is information on what products from our own and from competitor organizations are being used commercially, and the trade names mentioned in patent documents provide significant clues. We need to have them indexed, in context. Their roles must be identified, and trade names—which are frequently common words that occur countless times in any file—must be tagged as such. Some of this information is indeed captured today; this capture needs to be more complete, more systematic.

Roles, tags, context—these are vitally important concepts. I spent a number of years urging patent authorities not to waste more time and money developing full text systems for searching patents; I had remembered the disaster of Project Potomac, a costly and failed full text search program of the 1970s for US patents. I was wrong; full text patent databases have turned out to be invaluable, and I have certainly become a believer. Full text patent searching is a marvelous way of finding needles in haystacks. Full text searches of the latest week's material is especially effective, and can often be done without overly narrow search strategies, without producing unmanageable output. Nevertheless, full text isn't the answer to complex structure searches, and it has problems in dealing with context. Proximity operators help, but some concepts require massaging by skilled document analysts. Full text searches can result in the retrieval of unmanageable piles of low-relevance stuff, and I remain unconvinced that relevancy ranking search engines are the answer. If this applies to unenhanced full text in documents searchable through the traditional online techniques, how much more does it apply to some of the data soup out on the Internet?

I remain stubbornly convinced that we still need, and will continue to need, the ministrations of skilled document analysts. And those analysts must retrain themselves in what they are looking for. The claims alone aren't enough, and the examples alone aren't enough. They may or may not be enough when you're dealing with organic chemicals, whether simple or complex. Benzoic acid is benzoic acid, no matter how it was made, and the same will hold for a complex pharmaceutical compound. Gertrude Stein said that a rose is a rose is a rose, but polyethylene is not polyethylene is not polyethylene. In my main area of operation, polymer chemistry, the structure and therefore the properties of a plastic or rubber depend on how it was made. Ziegler polyethylene was enormously different from the low-density polyethylene developed in the 1930s, and I think that document analysts reflected that situation effectively once the Ziegler−Natta polymers arrived in the mid-1950s. But metallocene polyethylenes have properties very different from those of Ziegler polyethylenes, and I've spent much of the past decade trying to get the message through to our database producers that it is essential to convey the source of a polyolefin. The literature is full of patents with claims based on the properties of a polymer, patents in which the examples don't detail the preparation of the polymer but mention that it was made by a Ziegler or by a single-site catalyst. That mention frequently appears only in the front end of a patent specification, and it is absolutely nontrivial; it is vital information to those who are in the business, and it needs to be part of the document analysis. I urge all patent document analysts (and their supervisors) to understand that if the claims of a patent are based on the properties of a polymer or other material, it is imperative that document analysis reflects how those properties were achieved.

The same holds for polymerization process patents. I'm not breaking any new ground here, I've made the point more than once over the years, but polymers made with what is nominally the same catalyst system but in a slurry in an autoclave, or in a loop reactor, or in a gas phase fluidized bed system, or in a tubular flow reactor, are different. They have substantially different properties. Chemistry isn't everything; chemical engineering matters, too, and analysts must recognize that fact.

Some final observations on document analysis. If a patent's claims are based on a Markush formulation, it's not enough to cover that formulation, whose language can often be nearly impenetrable, without giving at least one concrete example. Markush or not, if the examples are far-ranging, it's not enough to cover just the first example without indicating what the scope is. And please, CAS, the thoroughness of your indexing is magnificent, but if you could figure out a way to build more context into the many substances that appear in some indexing records, so that we could tell what went with what, it would be greatly appreciated. A final request: it is more critical to be thorough when the original is in a relatively difficult language, such as Japanese. We can easily check out questions relating to English language documents; it's not that easy with a Japanese Kokai.

I hope that I'm not conveying the impression that I want the world turned upside-down to satisfy my particular needs. I do believe, though, that document analysts must learn to identify the truly significant information in technical documents, regardless of what long-standing rules or guidelines might say, regardless of long-standing habit. Russell Rowlett used to say that indexed abstracts provided a filter for the literature. If value-added processes are to survive and thrive, it is essential that this filter capture and transmit, accurately and clearly, that significant information.

The bottom line is quality. We depend on database producers for value-added products; our ability to serve our employers and clients rests on the tools we use. We should not have to second-guess those tools. They must be reliable. The world of scientific and technical information at the turn of this millennium is in a state of ferment, and many predict the disappearance of the system of databases that have become so much a part of our lives over the past 40 (and more) years. I believe that there is indeed a need for those information providers to continue to evolve, to develop new techniques and systems. If they do, if they remain committed to providing real value-added products, with high quality and timely delivery, I believe that they will continue to be an invaluable resource for chemical science and technology.

REFERENCES AND NOTES

(1) An additional 325 000 US patents were to issue by July 24, 2001.

(2) The first of these Pre-Grant Publications actually appeared on March 15, 2001.

(3) Girard, A.; Kaback, S. M.; Landsberg, K. APILIT and APIPAT: Petroleum Information Online. *Database* **1978**, *1* (2), 46−67.

(4) In 2001 the USPTO began to indicate some citations selected by the examiner, but chose to omit those citations also on the applicant's IDS, eliminating much of the potential benefit of identifying examiner citations.

(5) Kaback, S. M.: Online Patent Searching: The Realities. *Online* **1983**, *7* (4), 22−31.

(6) Kaback, S. M.: *WPI* + *APIPAT* The Sum of the Parts. *Online* **1990**, *14* (2), 22−26.

(7) Skolnik, H. Multiterm Index: A New Concept in Information Storage and Retrieval. *J. Chem. Doc.* **1970**, *10*, 81−84.

CI010088F