Enhanced access to information via personal bibliographic databases

Thomas E. Wolff

Amoco Research Center, Naperville, Illinois, USA

The Information Research and Analysis group provides scientific, technical and business information to customers within Amoco Corporation. In general, this information is retrieved from databases on the major online systems. Since members of our group have considerable and varied experience working in Research and Development within Amoco and elsewhere, we understand our customers' businesses and concerns and are well situated to carry out searches in accordance with their needs. We add value to our services by analysing the search results for relevance and editing the output for clarity and improved accessibility. Each report includes a cover letter describing search strategies used and results obtained. On request, additional technology assessment will be provided. However, our reports are generally collections of database citations with abstracts. The customer can then order the full documents as they deem necessary, doing his or her own assessment of the subject area.

We place considerable emphasis on providing search reports which can be most effectively used by our customers, and we look to add this value at minimal cost. Traditionally, our efforts have focused on printing reports with only the most appropriate amount of bibliographic or indexing information, or on organising reports into useful subject categories, occasionally adding a Table of Contents. But no matter how well the information is presented, reports are almost always used just once and just by the original requestor. It is clear to us that this is not an effective use of the company's resources. Information is too expensive to be used in such a short-term and short-sighted manner. Therefore, we have begun to provide our search results as bibliographic databases using selected personal computer programs. As will be discussed further, these databases provide our customers with a convenient means to evaluate the search results initially, to revisit these reports, and in some cases, to share that information with colleagues.

The development and usage of these bibliographic databases involves numerous process steps of varying complexity. As these are discussed, reference will be made to Figures 1 and 2 at the end of this article, which show patent and journal article citations, respectively. Each figure is divided into: (a) information downloaded from the online system; (b) personal computer 'screen' images of the information in a Library Master database; and (c) the citation information reformatted and output

from the Library Master database and automatically converted to a word processor document.

What are personal bibliographic databases?

Two years ago, the moniker Personal Bibliographic Databases seemed perfectly appropriate for the products we were developing. Individuals would request a search and the information was provided as a personal computer database along with a hard copy report if requested. So 'personal' referred to the personal requestor as well as to the PC. The PC was the medium of choice because no appropriate mainframe (IBM or VAX) database programs were available at Amoco and the purchase price for new ones was generally over $10,000, well beyond the means of any individual users. Now we find that research groups are requesting searches for the whole team to use. An individual's PC is too restrictive. We are again looking at mainframe programs, but more likely we will follow the distributed computing path and load the databases on local area networks. At that point, these may well be 'Team Bibliographic Databases' or 'Research Project Bibliographic Databases.'

The designation 'bibliographic' may not be appropriate in the future either. Right now we have difficulty incorporating chemical structures or numerical data, as from Beilstein, in our personal bibliographic databases. In addition, we look forward to adding graphics, either from CD-ROM records or from the online patent databases of the future. Finally, the addition of hypertext capability would further enhance the value of the customers' databases. We anticipate that software packages will improve considerably and require that any data imported into today's databases must be exportable in formats appropriate for tomorrow's improved programs.

The personal bibliographic databases of today are requested by individual researchers to augment or replace 'traditional' searches. Databases are particularly appropriate for broad searches of hundreds or thousands of citations or for periodically updated searches which gradually build up the database. Since the database software is well suited for citation categorisation and sorting, and for putting out reports in word processed formats, formation of the database can be a means to produce a better hardcopy report. In this case, the database may be a side benefit.

Search results from numerous databases and online systems are often combined into a single personal bibliographic database. Output from ORBIT, QUESTEL and STN is well-suited for importing because of the field-delimited formats with unique field tags and well-defined text formats. Information from Dialog Information Services may also be used, either in field-delimited 'tagged' or in native, untagged format. However, problems arise in either case. Tagged format is preferred, although one must deal with non-unique field designators and inconsistent formatting. Untagged information is probably best suited for importing into structureless personal databases, which are not further considered here, although it may be imported into structured databases using the special untagged-format Biblio-Links import/conversion program used with Pro-Cite bibliographic software.

Personal databases will soon be drawing upon other sources of information as well. Already, diskettes and CD-ROM are replacing many hardcopy sources, such as the Science Citation Index from the Institute for Scientific Information, which is

available on CD-ROM and includes abstracts.[1] The CABI (Commonwealth Agricultural Bureaux International) now provides 47 printed abstract journals on diskette.[2] Many 'full-text' journals are available online, such as the extensive Chemical Journals Online (CJO) on STN. These online journals suffer from lack of tabular and image information, which both the online services and personal bibliographic database software will have to address. Eventually, many journals will be published exclusively online. While this may seem unlikely, The Online Journal of Current Clinical Trials, a peer-reviewed journal, is the first of a series of anticipated publications from Primary Journals Online. Although many issues must be dealt with, including fees, copyright and distribution, a level of acceptance has been reached with BIOSIS now indexing articles from the Online Journal of Current Clinical Trials.[3]

The value of personal bibliographic databases

The use of personal databases may be characterised by the following: convenience, communication, idea generation and cost.

Convenience

Search results are in an accessible format with many of the search techniques available online and with enhanced sort and output capability. Information has extended usable lifetime, well beyond the 'read-once-and-file-away' lives of most hardcopy reports. Multiple searches may be combined into larger databases, or subsets may be conveniently separated into narrow-focus databases. The value of individual records may be enhanced by annotation, cross-referencing or combination of related records.

Communication

Databases can be used to create custom bibliographies and topical reports. Shared databases can be annotated by individual users to communicate key features of the cited references.

Idea generation

Convenient searching and browsing through information should lead to generation of new ideas. One's perspectives change as projects progress, so each time questions are asked and the data evaluated, there are new opportunities to learn and generate new search questions. When information in a personal database is found to be incomplete, or when relevant information cannot be found because of limitations in the database indexing or search software, new queries are made in the source online databases, which may then lead to an expanded personal database.

Cost

Some of the value in maintaining information in an accessible personal database may lead to actual cost savings. Certainly, many 'trivial' questions could then be answered on one's own PC. In addition, some queries of the source database may become 'unnecessary' because the answers have already been obtained in prior searches. On the other hand, the more 'information-aware' researchers should continue to generate more search questions as their knowledge of a field increases.

The net search costs may decrease or increase, but the savings in doing better research should be substantial when search information is well used.

Database packages

Many database software packages appropriate for developing personal bibliographic databases are available and have been reviewed [4-6]. Our experience is with IBM-compatible personal computers, for which we have evaluated four programs in detail:[7] Library Master; Notebook II; Papyrus; and Pro-Cite with Biblio-Links. Two of these have recently been upgraded, Library Master to a local area network version with improved functionality, and Pro-Cite to a more efficient version. Lately, we have also considered three other programs: EndNote, recently translated from the Macintosh version[8]; ideaList[9]; and STN Personal File System, essentially STN Messenger language for the PC with very efficient importing for downloaded STN records. On the basis of our evaluations and our customer acceptance, we have settled on Library Master and Pro-Cite as our recommended software. However, as stated earlier, we will be continually looking for improved ways to bring this information to our customers. For example, our department is evaluating document management software, such as Verity or Excalibur, for managing Amoco's internal documents. Many of these have sophisticated search capabilities. When Amoco moves from the mainframe environment and adopts one of these programs, we may find that they will also be well suited to our bibliographic databases.

Principal concerns regarding personal database development

User-definable Record Formats

The patent user community is not a significant constituent of the bibliographic software market. This is most clearly shown by the patent record type available with the software. Some programs list patent types as standard but the format is so limited as to be almost worthless. For example, Papyrus has a patent record type with only nine simple fields, and it is impossible to search or sort on assignee. In Pro-Cite, the patent record format, which is available as a supplemental workform, is acceptable in its simplicity and, even better, modifiable through standard Pro-Cite procedures. However, multiple paragraph abstracts, multiple entry fields (e.g., abstracts from multiple equivalents), and lists of applications or patent family equivalents cannot be easily imported or handled properly in Pro-Cite. These specially formatted fields are all 'word-wrapped' into single, continuous, difficult-to-read paragraph fields. Other database program producers have not considered patents at all. However, for some, the program's flexibility and power can enable creation of a useful patent record type.

The first two screens of a patent record as created by us Library Master in are shown in Figure 1b. These 'data input forms,' as the screens for editing and browsing are called, are fully user-customisable. All bibliographic information fields are shown on the first screen including the first eight lines of the abstracts. Multiple-element list fields, such as author, equivalent patent and applications, are fully viewable by moving the cursor to them, but only the first two lines are shown initially. The remainder of the abstract is also available through the expand-field option. The

second and subsequent screens in our Library Master patent records include a comment field, indexing, other abstracts, and patent claims. The first screen of a journal article record in Library Master (Figure 2b) shows the similarity we have maintained between the layouts of patent and article screens. For example, in each record type the first screen is headed by the title, followed by bibliographic information and then the abstract.

The patent record screen also shows new fields generated by parsing fields in the imported data or by modification of the downloaded information prior to importation. One useful conversion is the separation of patent number from its date, required since most patent files combine them in one information field. Another is the creation of the earliest priority number and date fields, information that is usually buried in other application fields in online records. These separate, searchable earliest priority application fields are valuable to simplify the identification of patents from the same family. We make substantial use of KEDIT macros to reorganise downloaded records, to make the information both more accessible in the personal bibliographic database and more consistent from source to source. For example, the information to be imported is reorganised to take advantage of the Library Master feature which allows any field to be designated as a date-format field. These fields can also have searchable text following the date. The application and patent number fields have dates first, which allows both dates and numbers to be searched appropriately.

Importing

As in most computer software applications, a balance must be struck between program flexibility and ease-of-use. The broad variety of information sources makes critical the program's flexibility in importing process. At the same time, this process should be straightforward and efficient, since users generally wish to pass quickly through the importing stage and on to using the information. Software producers have met the importing challenge various ways. Most create separate importing modules or programs. For our purposes, EndLink, the importing program for EndNote, errs on the side of simplicity because it is a 'black box,' as described by a sales representative, which allows for no alteration of the downloaded information. At the other end of the scale, Papyrus requires development of a complicated import template which must account for all possible variations in the imported information. Although the Papyrus producer will create customised importing programs for the users, we have found them to be generally ineffective, as for importing Chemical Abstracts information from STN, for example. Developing import templates can be much more straightforward. The Convert program with Notebook II is extremely simple to customise, and the import facility for Library Master is nearly as clear-cut. The balance is probably best struck by Pro-Cite, which provides separate Biblio-Links for each online systems and many CD-ROM products. The Biblio-Links can also be modified readily. However, until recently, the pricing policy for the Biblio-Links made them considerably more expensive than the database software Pro-Cite itself; with the newly created package sets, the import and database software are more nearly matched in cost.

Responsibility for importing difficulties also lies with the database producers. Personal computer importing programs must be powerful because online information is so complex and inconsistent. Some databases have dozens of document types. For example, at last count, Compendex has 26 document types on STN. Of more concern, fields frequently contain more than one type of information, necessitating the use of sophisticated field parsing or editor programs. For example, in APILIT, the source field has almost no consistent format at all and can contain CODEN or ISSN numbers or references to other sources such as Petroleum Abstracts. Similarly, the source field in CA file article citations contain all the bibliographic information except year of publication (see Figure 2a). CA file patent record source (SO) field does not contain the application or patent numbers or issue date, as might be expected, but rather the 'country' information, number of pages, and, for Polish patents only, a phrase which describes that the abstract was taken from the application. Rather, the actual 'patent information' is found in the PI field.

Online database information can also be presented badly. For example, the tagged formats for Dialog Information Services appear to be just an afterthought created out of the 'normal' format. Dialog field tags frequently are not unique, and they may contain subfield descriptors in angle brackets at the beginning of the field's text area (see Figure 1a of a Derwent World Patent Index record). Dialog also uses the virgule or vertical bar to designate end-of-field, but this mark is not applied consistently throughout records or between files. Our KEDIT macros make major changes to Dialog tagged output.[7] In addition, we have recently written macros which convert untagged Derwent output (e.g., Dialog format 7) to the equivalent tagged format. This allows us to use Derwent citations for importing, even if we had not planned on it when the search was originally downloaded. The whole issue of having to remember to use tagged format does not exist when downloading information from ORBIT, Questel, or STN. These online systems only produce tagged output, which is generally more consistent than that from Dialog.

The customer-oriented approach of the CABI seems particularly enlightened.[2] Subscribers to CAB Abstract Journals on diskette requested search and retrieval software. Other database providers have responded by including specialised retrieval software with their data, ISI (Science Citation Index on diskette) or Ziff-Davis (Computer Library on CD-ROM). However, CABI recognised that subscribers had their own preferred database software which could be used for information from varied sources. To accommodate this need of the majority of their subscribers, CABI publishes their abstracts journals in both comma-delimited and Pro-Cite proprietary formats and provides information on potentially useful database software. The lesson to be drawn is that information should be formatted consistently and thoughtfully in consideration of importation into the personal databases of the customer's choice.

Browsing, searching and general ease of use

The forecasted demise of hardcopy information is often regretted because of the value and pleasure in browsing and the serendipity in finding unsought-for information of interest. This ability to browse information conveniently is critical to idea

generation but need not necessarily be lost with computer databases. But software designers must remove barriers to information access. For example, consider the situation in which a search has been carried out in a personal database and the user wishes to browse the results. One of three approaches is taken by software producers. The first is to provide just enough information to enable a judgment about whether the 'hit' citation should be further reviewed. This information might be a page of author and title information. The problem with this implementation is that it generally takes two or more keystrokes to move from abstract to abstract, whether the user looks at all retrieved records or selects on the basis of the limited scan information. The second approach is to have the user move from citation to citation. However, if the screens are not well designed, one may again need multiple keystrokes to go from abstract to abstract. A third approach would be browsing actual occurrences of hit terms wherever they occur in the 'hit' record. This approach is available in some document management software (e.g., Magellan or Verity) but not any on the personal database software evaluated by us.

The ability to browse in a database generally seems to parallel other functions of the database software. That is, if it is straightforward to get around the data, it is often easy to get around the software. Most of the programs we have evaluated have menu systems and short-cut keystrokes to skip through, but not avoid, the menus. With few exceptions, our customers would prefer good menu-driven programs to command-line programs. However, with experience, menus can become tedious. The ease of database searching is another tell-tale area. Boolean operators are generally available, but each software program has its own rules about operator order, use of parenthesis, and other syntax, e.g., the use of 'not' vs. 'and not' vs. the caret symbol (^). Similarly, truncation operators are generally straightforward, although the specific characters used vary considerably. Proximity operators are generally not available. Also of concern is the layout of the search information screen, the format for entering field codes (e.g., from a menu of useful abbreviations vs. unrecognisable field code numbers) and the convenience of accessible lists of database index terms for searching. Other database functions, such as editing, sorting and set handling, frequently follow conventions already used in the browse and search capabilities.

Report generation

Database software should integrate well with other personal computer operations, especially for reports generated for importing into word-processing programs. Most personal bibliographic database software is designed to produce bibliographies for incorporation into published works, and most leading word processing programs have been accommodated. In some cases, the database program can also scan an article in word-processor format, find reference indicators and create properly ordered and formatted reference lists. The database programs may have options to conform with dozens of bibliography syntaxes required by journals, which relieves the author or secretary from much tedious footnote reformatting.

For our work as industrial scientists and engineers, we need to create subject bibliographies or reports based on online search results. The output syntax is seldom

specified by the recipient, who only needs the information complete and in easily-readable format. Therefore, we have created custom output formats which resemble the 'CBIB AB' display format of the CA file on STN (see Figure 1c and 2c). Our output formats always have the title first, in bold-face, followed by pertinent bibliographic information, some of which is italicised for ready identification, then lists of equivalent patents when available, and the abstract(s). When the basic patent is non-US but a US patent abstract is available, both the basic and US-equivalent abstracts will be printed. Index terms are rarely printed, except when the 'hit' terms, such as Registry numbers, would be useful to the recipient. This bibliographic information is usually output in word-processor format for generation of readable hardcopy reports. However, for some of our reports which are viewed or transferred electronically, ASCII text is preferred.

Cost

One of the first concerns of the potential customer is about the cost of the personal bibliographic software program. After all, they think, if the software is expensive, there is no point in considering further the development of a personal database. Even after hearing that the database program cost is a very small fraction of the total cost of generating the database itself, customers still want to purchase inexpensive programs to keep their initial investment down. Fortunately, the effective personal computer database software is inexpensive, costing at most a few hundred U.S. dollars. The STN Personal File System is one of the more expensive personal computer programs at $600, which is enough for most of our customers to look elsewhere. But this price is still very reasonable compared to mainframe programs, which may cost $10,000 to $20,000 and more depending on the computer platform. The exception is Papyrus, whose VAX-based version is equivalent in cost ($99) and functionality to the personal computer edition, at least up to version 6.0. This low cost has attracted several of our customers, but they have never been satisfied with the program for reasons described previously.[7] The high cost mainframe (VAX and IBM) programs are completely out of the reach of our employees, until the usage of personal databases reaches the critical number required to justify the software purchase. By this time our use of the PC-based programs will probably be entrenched.

As already suggested, the cost of developing personal bibliographic databases involves far more than the software purchase. The principal costs are: downloading information, or typing it in manually; editing or reformatting the downloaded information; importing the information into the personal database; and copyright or licensing fees. The most expensive portion is for information downloading, which includes search strategy development, online time and citation print charges. The cost per record ranges generally between two and five dollars, depending on database and online service. Citation reformatting and importing are carried out as automatically as possible using editor macro programs and standardised import routines.[10] The incremental cost for reformatting and importing is kept to a minimum, but can still run at one to two dollars per record. Finally, licensing fees are frequently required for storing the information electronically, as will be discussed later. These fees can be another one dollar per record per year. Thus, the

total cost of personal bibliographic database development seems to average about $6 per record, or thousands of dollars for hundreds of records. As stated earlier, the cost for the personal database software itself is a hardly significant fraction of the total personal database development cost.

Local area network (LAN) access

Convenient sharing of information among group or team members will be a critical requirement of bibliographic databases in the future. Our customers are thinking about this cost-effective approach to information usage, but both the hardware and software are holding us back. Many of our employees are still working on mainframes, and the conversion to personal computers is taking place slowly. Those fortunate individuals with personal computers are not yet connected to their own local area networks. We are considering loading our customers' bibliographic databases on our LAN, at least until their own networks are established. For this, we would use Library Master, whose LAN version is undergoing extensive beta-testing prior to release in late summer. Multiple users have full, simultaneous access to any database, except that no two people can have simultaneous write access to the same record in a database. We have been very impressed with the implementation of the LAN version of Library Master. Since the licensing cost per user is comparable to the cost of individual copies, we anticipate having our whole group using the LAN version. This will facilitate purchase and installation of database software upgrades.

Non-textual information

Bibliographic database software will need to accommodate more than simple ASCII textual information, as our sources become more complicated. Even importing chemical structural information in text mode (as opposed to the graphics mode available with 'type 3' terminals) from the Registry file on STN cannot generally be imported without 'word-wrapping' the characters into meaningless strings. Other text information frequently handled improperly include lists, such as patent equivalents, tables as output with the Dialog report format, or numerical data. The software packages that we currently use have no means of incorporating images in binary formats. As images become more prevalent on CD-ROMs or online, e.g. Registry file structures or the online Derwent abstracts of the future, we will probably want some hypertext pointers in our databases.

Other significant issues

Comprehensiveness

Personal bibliographic databases are not comprehensive. This is an important principal for all to understand. Search professionals often search multiple sources to increase the likelihood that 'all relevant" information is retrieved. However, cost-benefit considerations usually prevent searching every conceivable source. Even if the imported information is 'complete,' the personal bibliographic database must be continually updated as the source files are. Furthermore, all search techniques available online are not available on personal computers, e.g. proximity operators or online database coding. So, 'relevant' citations in the personal database

may be missed on subsequent searching. Personal databases are valuable information repositories which may reduce, but can never replace, searching the source literature and online files.

Copyright, licensing agreements, and cost

Most of the information we use in our personal bibliographic databases is copyrighted, and appropriate licensing agreements have been established with the information producers. It is easy to empathise with the headline that copyright and licensing agreements are 'Making Life Complicated.'[11] This results from the conflicting views of information users and producers. The former tend to think of information as free, while the latter require fair compensation for creating and assembling the information. We search intermediaries are caught in the middle, with particularly strong influence from corporate policy requiring full compliance with the law. However, the latest major revision of American copyright law, the Copyright Act of 1976, "was not designed with information brokers [or] rapid computer transmission of information . . . in mind."[12] Personal computers were not even invented then, so what do we do about sharing electronic information on local area networks? The answer seems to be that we work out individual licensing agreements with every information provider. Even having source files on the same online system is no benefit, because the online systems are just conduits of the information and the licensing policies. In the near term, the best goal to aim for might be having similarly structured licensing policies which would differ in cost but not billing structure.

The factors influencing database storage licensing costs should be: (1) lost revenue from reduced online searching or journal subscriptions, and (2) recovering fees for multiple distribution of information. Information producers also suggest that information users should pay for the convenience of electronic storage of information, although I believe this is a red herring. One should not be penalised for electronic storage for information when hard copy storage is free. However, if electronic storage means reduced revenue for the producer, then the producer may be entitled to make up this revenue via licensing fees. Similarly, fees are paid for photocopying hard copy materials via the Copyright Clearance Center in the United States, but there is no parallel organisation or payment mechanism for copying electronic information. Note that the difficulty of monitoring what documents are photocopied is magnified when translated to monitoring electronic access of shared database material or electronic distribution beyond the initial group of recipients. For example, very few database programs have any sort of access monitoring or copy-protection schemes. Some producers have suggested that any fees be paid on the number of potential information users, i.e., all personnel connected to a local area network, rather than on 'actual' users. In one such case, a licensing fee algorithm for 5000 citations available to 8-10 people calculated a fee of about $15,000 per year. After we got this proposal we found that there might be 50 interested users. Even the information provider realised that the algorithm broke down as the whole database could be licensed for $25,000 per year.

The most manageable licensing fee schedule would be based on the amount of information, i.e., the number of database records, rather than on the number of users with either potential or actual access. The latter requirement is unworkable and untenable with today's software. Furthermore, we already spend considerable time just complying with the simpler licensing fees based on number of records. We would prefer agreements which accept annual payment based on actual records in personal databases at the beginning of the agreement year. Additional records added during the course of the year would be exempt from payment until the next year. Some of our current licensing agreements have this structure.

To ensure that proper copyright credit is given in personal bibliographic database records, database producers and online systems should have the copyright statement as an integral part of each record. Copyright information is frequently found in the header information when the database is first entered or as a separate line or field before or within each record. This information is seldom incorporated into the personal database or subsequent output. We use KEDIT macros to append copyright statements to accession numbers, which are always printed as part of the customised record format (see Figures 1c and 2c).

Conclusions

- Our customers will continue to look for cost-effective ways of handling information, and we need to be at the forefront of this issue.
- Personal bibliographic databases will gain in acceptance as long as the cost of developing and maintaining them is reasonable compared to producing hard-copy searches.
- Customers are willing to pay a 'fair' price for access to information in their personal databases. Producers will have to work with their customers to determine this fair price, or they will lose this portion of their market.
- Electronic storage licensing fees should be based on the amount of information and not require an exact counting of users of that information. Fee scheduling over broad numerical ranges of users would be tenable.
- Straight bibliographic database software will satisfy most of our customers' needs in the near term. However, as the computerised information environment evolves, so too will the database programs of choice. Therefore, it is critical that we be able to export data from the currently-preferred software if that software does not continue to meet customer needs.
- Database producers and online systems must appreciate that searchers are using advanced computer technology to obtain and process downloaded information. The format of that data should be improved to allow everyone to readily take advantage of the technology. Specifically, database fields should be rationalised and simplified so each contains only one data element. Format consistency should be a hallmark of the database.
- Personal software developers need to continue to improve the process of browsing information so that users will feel as comfortable just 'cruising'

through the database as they would a hardcopy reference. Similarly, searching should be improved, at minimum by avoiding symbols where words should work, by adding proximity searching and possibly by including non-Boolean search techniques such as frequency searching or relevance ranking, or hypertext, techniques often found in document management programs.

References

[1] Garfield, Eugene. 'New Chemistry Citation Index on CD-ROM Comes with Abstracts, Related Records, and KeyWords Plus.' *Current Contents*, **No. 3** (January 20, 1992): pp. 5-9.

[2] Powell, Andrea. 'Making the Most of an Electronic Journal.' C.A.B. *International Database News*, **No. 10** (Nov. 1991): p. 2.

[3] 'BIOSIS to Index Articles from The Online Journal of Current Clinical Trials.' *Database* **15**, *No. 3* (June 1992): p. 12.

[4] Lundeen, Gerald W. 'Software for Managing Personal Files.' *Database* **13**, *No. 3* (June 1989): pp. 36-48.

[5] Lundeen, Gerald W. 'Bibliographic Software Update.' *Database* **14**, *No. 6* (Dec. 1991): pp. 57-67.

[6] Stigleman, Sue. 'Bibliography Formatting Software: A Buying Guide.' *Database* **15**, *No. 1* (Feb. 1992): pp. 15-27.

[7] Wolff, Thomas E. 'Personal Bibliographic Databases: An Industrial Scientist's Perspective.' *Database* **15**, *No. 2* (April 1992): pp. 34-40.

[8] Mead, Thomas. 'Making the Link: Importing Downloaded Bibliographic References to Pro-Cite and EndNote on the Macintosh.' *Database* **14**, *No. 1* (Feb. 1991): pp 35-41.

[9] Stigleman, Sue. 'ideaList: Flexible Text Storage and Retrieval Software.' *Database* **15**, *No. 3* (June 1992): pp. 50-56.

[10] Wolff, Thomas E. 'KEDIT: Text Editor for Post-processing Searches.' *Database* **15**, *No. 3* (June 1992): pp. 43-49.

[11] Collier, Harry (Ed.). 'Making Life Complicated.' *Library Monitor* **33** (July 1991): pp. 1-3.

[12] Berring, Robert C. 'Copyright: Online Abstracts Raise New Issue.' *Database* **11**, *No. 3* (June 1988): pp. 6-9.

Information on packages

Library Master version 1.3: Balboa Software, 61 Lorraine Drive, Willowdale, Ontario M2N 2E3 Canada (416/730-8980; e-mail address hahne@epas.utoronto.ca), $199.95, $599 for a 5-node Local Area Network version, $5.00 for a demo version.

Pro-Cite version 2.0 for IBM-compatible PC or Macintosh with Biblio-Links: Personal Bibliographic Software, Inc., P.O. Box 4250, Ann Arbor, MI 48106 (313/996-1580), Pro-Cite $395, Biblio-Links at various prices.

KEDIT version 5.0: Mansfield Software Group, Inc., P.O. Box 532, Storrs, CT 06268 (203/429-8402), $150 (DOS only), $175 (DOS and OS/2), free demo version. Technical support on MSG bulletin board at 203/429-3784; CompuServe (GO PCVEN); BIX (JOIN MANSFIELD). KEDIT macros, including many alluded to in this paper, may be found in the PCVEN library on CompuServe and on the MSG bulletin board.

43/4/24
AX- 91-193135/26
AX- <XRAM> C91-083597
TI- Prodn. of aromatic anhydride(s) and ester(s) with superior colour — by treatment of crude prod. with activated boric acid and fractionation, useful in polymer prodn.
PA- (STAD)_AMOCO CORP
AU- <INVENTORS> PARK C M; COATES R; HOLZHAUER J K; PETERSON J VI
NP- 002
PN- <BASIC> WO 9108204_A_910613_9126
PN- <EQUIVALENTS> EP 455802 _A_911113_9146
AN- <PRIORITIES> US 606603 (901031; US 443564 (891129)
AN- <APPLICATIONS> WO 9OUS6938 (901128); EP 90901071 (901128)
LA- English
CT- US 2971011; DE 1948374; US 3888921; US 4794195; 3 Jnl.REF
DS- <NATIONAL> JP
DS- <REGIONAL> AT; BE; CH; DE; DK; ES; FR; GB; GR; IT; LU; NL; SE
AB- <BASIC> WO 9108204
Prodn. of aromatic anhydrides and esters with improved colour comprises treatment with activated boric acid followed by fractionation at 200—275 deg C and 25-1 mmHg. The boric acid is activated by heating with an organic hydrocarbon acid or anhydride.
Specifically claimed is treatment of trimellitic anhydride (TMA) and dimethyl-2,6-naphthalene dicarboxylate.
USE/ADVANTAGE — TMA is used an an intermediate in the prodn. of quality plasticisers and ployester resins.
Dimethyl-2,6-naphthalene dicarboxylate is a monomer used in the prepn. of high performance polyesers, esp.
poly(ethylene-2,6-naphthalene) (PEN) which is used in 'hot- filled' food and beverage containers, tyre cord and magnetic recording tape. The delta E colour of TMA is improved from, e.g. 2.69-0.44 in a process which does not require expensive recovery and regeneration of dehydration agents. @22pp Dwg.No.0.0)@
FS- CPI
DC- A41; E13; E14;
IC- C07C-051/42; C07C-067/48; C07C-069/76; C07D-307/77
MC- A01-E11; A08-P03; E06-A02A; E10-G02A
DR- 1894-U; 1924-U

Fig. 1a: Derwent World Patent Index (Dialog) citation as downloaded in tagged format

BROWSE DB: 1 AMOCO ID#: 1571 REC: 20 SC: 1

RECORD TYPE: PATENT

TITLE:	Prodn. of aromatic anhydride(s) and ester(s) with superior colour - by treatment of crude prod. with activated boric acid and fractionation, useful in polymer prodn.		
AUTHORS:	PARK, C M COATES, R	ASSIGNEE:	(STAD) AMOCO CORP
PATENT NUM:	WO 9108204 A	DATE:	13 Jun 1991
APPL NUMBER:	28 Nov 1990 EP 90901071		
APPL PRIORIT	US 443564	DATE PRIORIT:	29 Nov 1989
EQUIV PATENT	13 Nov 1991 EP 455802 A 31 Mar 1992 US 5101050 A	ALL PRIORITY:	29 Nov 1989 US 443564 31 Oct 1990 US 606603
ACCESS. NUM:	91-193135/26 - Copyright	LANGUAGE:	English
USE:	3 TYPE: 9202 92-	LOCATION:	

ABSTRACT:

BASIC: WO 9108204
Prodn. of aromatic anhydrides and esters with improved colour comprises treatment with activated boric acid followed by fractionation at 200-275 deg C and 25-1 mmHg. The boric acid is activated by heating with an organic hydrocarbon acid or anhydride.
Specifically claimed is treatment of trimellitic anhydride (TMA) and dimethyl-2,6-naphthalene dicarboxylate.
USE/ADVANTAGE - TMA is used as an intermediate in the prodn. of quality

1HELP 2PRVIEW 3OLD CONT 4DUMP 5PRINT 6BRIEF 7END 9EXPAND PGDN NXT PGUP PRV ESC QUIT

Fig. 1b: Derwent WPI citation as a Library Master database record (Screen 1)

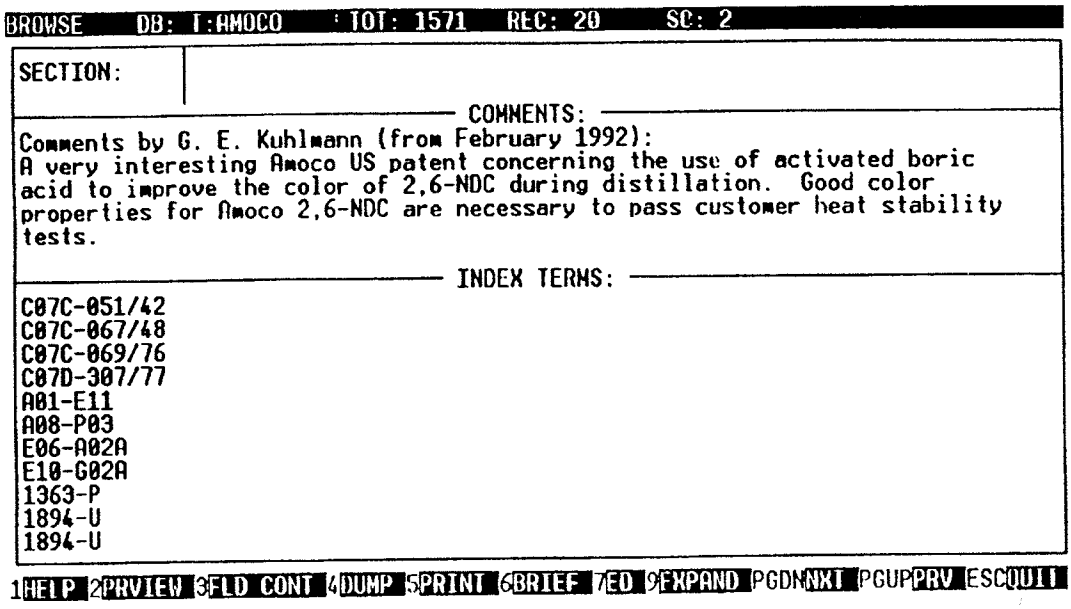

Fig. 1b: Derwent WPI citation as a Library Master database record (Screen 2)

Prodn. of aromatic anhydride(s) and ester(s) with superior colour — by treatment of crude prod. with activated boric acid and factionation, useful in polymer prodn.

Park, C.M.; Coates, R.; Holzhauer, J.K.; Peterson, J.V. ((STAD) AMOCO CORP). *WO 9108204 A*, 13 Jun 1991 (Appl. US 443564, 29 Nov 1989) (English) [91–193135/26 — Copyright (C) 1992 Derwent Publications Ltd. — For Internal Use Only].

BASIC: WO 9108204

Prodn. of aromatic anhydrides and esters with improved colour comprises treatment with activated boric acid and followed by fractionation at 200–275 deg C and 25–1 mmHg. The boric acid is activated by heating with an organic hydrocarbon acid or anhydride.

Specifically claimed is treatment of trimellitic anhydride (TMA) and dimethyl–2,6–naphthalene dicarboxylate.

USE/ADVANTAGE — TMA is used as an intermediate in the prodn. of quality plasticisers and polyester resins. Dimethyl–2,6–naphthalene dicarboxylate is a monomer used in the prepn. of high performance polyesers, esp. poly(ethylene–2,6–naphthalene) (PEN) which is used in 'hot- filled' food and beverage containers, tyre cord and magnetic recording tape. The delta E colour of TMA is improved from, e.g. 2.69–0. 44 in a process which does not require expensive recovery and regeneration of dehydration agents.

@(22pp Dwg.No.0.0)@

Comments by G.E. Kuhlmann:

A very interesting Amoco US patent concerning the use of activated boric acid to improve the colour of 2,6–NDC during distillation. Good colour properties for Amoco 2,6–NDC are necessary to pass customer heat stability tests.

Fig. 1c: Derwent WPI citation as output by Library Master in Word Perfect 5.1 format

L35 ANSWER 31 OF 45 COPYRIGHT 1992 ACS
AN CA116(12):108732t
TI Deactivation mechanisms in liquid phase oxidations caused by carboxylic acids
AU Partenheimer, W.; Kaduk, J. A.
CS Amoco Chem. Co.
LO Naperville, IL 60566, USA
SO Stud. Surf. Sci. Catal., 66(Dioxygen Act. Homogeneous Catal. Oxid.), 613–21
SC 45–4 (Industrial Organic Chemicals, Leather, Fats, and Waxes)
SX 25
DT J
CO SSCTDM
IS 0167–2991
PY 1991
LA Eng
AB The addn. of selected arom. acids to the Co/Mn/Br-catalyzed, homogeneous liq.-phase autoxidn. of 1,2,4–pseudocumene in aq. HOAc reveals two different types of catalyst deactivation. One type occurs without catalyst metals pptn. and it is suggested that this is caused by Co(III) decarboxylation of the arom. acids. The second type is caused by the decrease of the catalyst concn. due to the pptn. of the metals as their arom. acid compds. The x-ray crystal structures of Co and Mn pyromellitate suggest that the driving force for the pptn. may be H bond formation.
KW pseudocumene oxidn catalyst deactivation; arom acid deactivation oxidn catalyst
IT Hydrogen bond
(in manganese pyromellitate, autoxidn. catalyst deactivation in relation to)
IT Oxidation, aut-
(liq.-phase, of pseudocumene, in presence of cobalt acetate-manganese acetate-sodium bromide, effect of arom. acids on)
IT Crystal structure
(of manganese pyromellitate, deactivation of pseudocumene oxidn. catalysts in relation to)
... ...
IT 7647–15–6, Sodium bromide, uses
(catalysts contg., for oxidn. of pseudocumene, deactivation of, mechanism of)
IT 71–48–7, Cobalt diacetate 638–38–0, Manganese diacetate (catalysts, for oxidn. of pseudocumene, deactivation of, mechanism of)
IT 56004–36–5
(crystal structure of, deactivation of pseudocumene oxidn. catalysts in relation to)

Fig. 2a: Chemical Abstracts file (STN) journal citation downloaded

BROWSE DB: 1:AMOCO IOI: 1550 REC: 14 SC: 1

RECORD TYPE: JOURNAL ARTICLE

TITLE:	Deactivation mechanisms in liquid phase oxidations caused by carboxylic acids		
AUTHORS:	Partenheimer, W. Kaduk, J. A.	ASSIGNEE:	Amoco Chem. Co.
TRANS TITLE:			
PERIODICAL:	Stud. Surf. Sci. Catal., 66(Dioxygen Act. Homogeneous Catal. Oxid.), 613-21		
VOLUME:	ISSUE:	DATE:	1991
PAGES:		ABBREVIATION:	
ACCESS. NUM:	CA116(12):108732t - Copyr	LANGUAGE:	English
USE:	11 TYPE: 92-0574	LOCATION:	Naperville, IL 60566, USA

ABSTRACT:

The addn. of selected arom. acids to the Co/Mn/Br-catalyzed, homogeneous liq. -phase autoxidn. of 1,2,4-pseudocumene in aq. HOAc reveals two different types of catalyst deactivation. One type occurs without catalyst metals pptn. and it is suggested that this is caused by Co(III) decarboxylation of the arom. acids. The second type is caused by the decrease of the catalyst concn. due to the pptn. of the metals as their arom. acid compds. The x-ray crystal structures of Co and Mn pyromellitate suggest that the driving force for the pptn. may be H bond formation.

1HELP 2PRVIEW 3FLD CONT 4DUMP 5PRINT 6BRIEF 7FD 9EXPAND PGDNNXT PGUPPRV ESCQUIT

Fig. 2b: Chemical Abstracts file (STN) journal citation as a Library Master database recorrd (Screen 1)

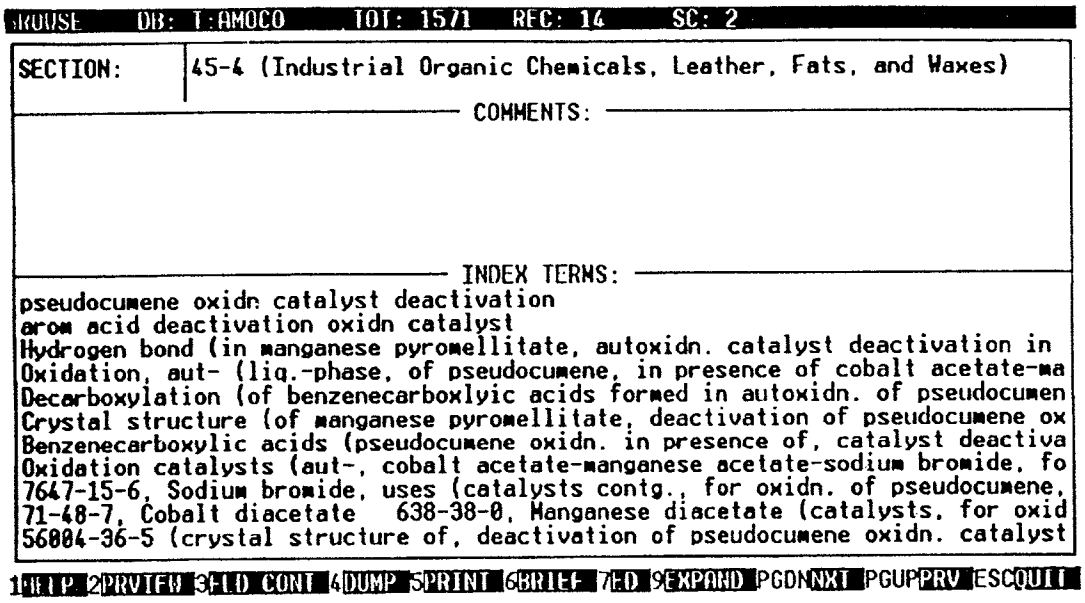

Fig. 2 b: Chemical Abstracts file (STN) journal citation as a Library Master database record (Screen 2)

Deactivation mechanisms in liquid phase oxidations caused by carboxylic acids

Partenheimer, W.; Kaduk, J.A. (Amoco Chem. Co., Naperville, IL 60566, USA) *Stud. Surf. Sci. Catal., 66(Dioxygen Act. Homogeneous Catal. Oxid.) 613–21* (1991) (English) [CA116(12):108732t—Copyright (C) 1992 ACS].

The addn. of selected arom. acids to the Co/Mn/Br-catalyzed, homogeneous liq.-phase autoxidn. of 1,2,4–pseudocumene in aq. HOAc reveals two different types of catalyst deactivation. One type occurs without catalyst metals pptn. and it is suggested that this is caused by Co(III) decarboxylation of the arom. acids. The second type is caused by the decrease of the catalyst concn. due to the pptn. of the metals as their arom. acid compds. The x-ray crystal structures of Co and Mn pyromellitate suggest that the driving force for the pptn. may be H bond formation.

Fig. 2c: Chemical Abstracts file (STN) journal citation as output by Library Master in Word Perfect 5.1 format

Harry Collier (Editor)

Proceedings of the Montreux 1992 International Chemical Information Conference

Annecy, France 19–21 October 1992