Information to order

An information research and analysis team can be a critical component of R&D by developing and maintaining customized current awareness products.

Robert E. Buntrock
Thomas E. Wolff

We live in a time of change, and the rate of change keeps increasing (*1*). This is especially true for scientists and technologists. For years, even decades, the half-life of knowledge of the typical graduate engineer has been estimated at about 5 years (*2*). Even in the general populace, more people perceive that education is not static and that almost everyone must continue to be educated, trained, and informed throughout his or her lifetime. Being informed produces a plethora of pronouncements ranging from the dawn of the "information age" to the brain death of those poor souls inundated by the "information explosion." The difficult task of gaining control of information has been described in these pages before (*3*).

At the Amoco Research Center, scientists and engineers can work with the Information Research and Analysis (IR&A) team to control the information they require. The team approaches potential information overload proactively by developing and maintaining customized current-awareness products. Maximizing the value of our services requires considerable versatility and ingenuity to meet the many and conflicting needs, most notably low cost and appropriate subject coverage. Here we will describe some of the methods we use to help our customers acquire current technical and business information. We will also discuss our requirements-gathering process and some specific products that we offer our customers.

The authors are with Amoco's Information Research and Analysis team.

Our emphasis will be on chemistry and chemical technology because we are information providers in a large petroleum and petrochemical company. Not only is chemistry the "central science" (*4*), but chemical information has unique aspects not found in other forms of information. Any effective information system for chemistry and chemical technology must address chemical structures, reactions, and processes. In addition, the ubiquitous issues of textual, including full-text searching, and numeric information are factors to be addressed.

Who are we?

The IR&A team is part of the Amoco Corporation Technology Support Department, which also includes analytical, statistics, computer, and process instrumentation services and other similar support groups. The IR&A team is a definite part of the research process at Amoco. We have not been directly associated with the library for more than four years, although we maintain offices near the central and satellite libraries and work with the library staff to satisfy mutual customers.

The five members of IR&A—Robert Buntrock, Susan Kaley, Alan Stewart, Robert Wilke, and Thomas Wolff—are chemical industry specialists with degrees in chemistry and engineering and years of experience "at the bench" at Amoco and other chemical companies. That experience translates into empathy for our customers and ready understanding of

0009-2703/94 $4.50

their concerns and needs.

The information skills of the group cover technical literature, patents, business and marketing information, and competitive intelligence. As part of Amoco Corporation Information and Computer Services in Naperville, IR&A members take advantage of the latest in computer technology both to retrieve information and to deliver it to our customers. The reports they provide come in many forms: organized hard-copy documents, e-mail notes, or searchable and browsable personal bibliographic databases.

How am I going to keep up?

Acquiring information is a fundamental human process, something so instinctive that many of us don't think about it. Top-notch scientists and technologists consider "doing research" a matter of acquiring, processing, and creating information. Taken for granted is a personal reading program involving many relevant journals, newsletters, and multiclient studies. Our brochure on current awareness for our corporate customers begins, "Nothing beats a good reading program to keep up with developments in your area of specialty or interest. However, even good reading programs often need to be supplemented . . ." (5).

But even the traditional current-awareness methods are undergoing rapid change. Electronic journals are beginning to appear and electronic bulletin boards are proliferating rapidly over the scientific landscape. Traditional newsletters and multiclient studies are evolving into various electronic species, disseminated by online systems or electronic mail, or on computer diskette or CD-ROM. Although many of our customers use these products and evolve with them, they still favor personalized, customized current awareness services targeted to critical research areas.

The goal of the information provider is to maximize the value of services while minimizing customers' costs. Because the Information Research and Analysis team charges for services on a complete cost-recovery basis, the cost of our services is highly dependent on customization and the amount of specialist activity. We consider the factors shown in Table 1 when determining the most appropriate medium for our customers.

Database selection follows directly from subject definition. For chemistry-related current awareness, an important issue is the type of information required, particularly whether patent information is central to the subject. Table 2 shows exemplary databases fulfilling needs for patent and non-patent information. Of course, some databases contain both.

Many databases offer current-awareness topic bulletins, which are often cost-effective because of the economies associated with having many subscribers. Chemical Abstracts has a large catalog of CA Selects bulletins produced biweekly from each update of the CA File. A newer product with fewer titles is the CASurveyor. CD-ROMs updated quarterly with the previous year's Chemical Abstracts records on the title subject. Derwent publishes Customized Bulletins on many specific technologies. The American Petroleum Institute has similar products of interest to the petroleum and petrochemical industries. Other products with bibliographic and abstract information may also be of interest, such as Drexel Polymer Notes from Drexel University.

We find that most commercial topic bulletins inadequately match researchers' needs, either because of

Table 1. **Factors to consider when designing a current-awareness product**

Subject	technical literature, patents, and/or industry news
Overall scope	emphasis on precision by including only the most relevant records, or on comprehension (completeness) by covering the interest area broadly
Sources	single or multiple information sources, whose information may be combined or kept separate
Cost	overall or based on each recipient
Number of recipients	the individual researcher, a project team, an entire product or business area, the whole subsidiary
Customer contributions	optional value-added comments or other contributions from customers for sharing with multiple recipients
Update frequency	either the default update period of proposed source or a time period more convenient to the customer(s)
Report organization	in natural source output order or categorized for customer comprehension and convenience
Record formatting	as downloaded or reformatted by the information specialist, involving simple changes or extensive editing
Duplicate removal	most important when using multiple sources, but valuable for some broad coverage information sources as well
Delivery medium	hard copy reports, electronic text or word processor files, groupware records, or electronic databases
Delivery mechanism	standard mail delivery, perhaps multiple copies or single copies with subsequent internal distribution, or electronic distribution via mainframes, distributed computing networks, or internal or Internet electronic mail
Archival access	future review made possible by electronic storage as text files or databases

Table 2. **Some sources of chemistry-related information**

Patents	Non-patent literature	Business/trade press
Chemical Abstracts (CA)	Chemical Abstracts (CA)	Chemical Industry Notes
CA Previews	CA Previews	PROMT
API Patent (APIPAT)	API Literature (APILIT)	API Business (APIBIZ)
Derwent World Patents Index	RAPRA	COMLINE
IFI/Plenum U.S. CLAIMS		PLASNEWS
USP U.S. Patents		
JAPIO		
EPAT/PCTPAT		

Sources of online chemical information

The Chemical Abstracts (CA) file online is updated on chemical topics with more than 500,000 records yearly from over 12,000 sources worldwide, including patents, journals, and conference proceedings. The file contains powerful indexing and Chemical Abstracts Service (CAS) Registry Numbers in addition to the bibliographic information on all online systems (STN, Dialog, ORBIT, Questel), but contains the CAS-copyrighted abstracts only on STN. Early access to CA records is available in the **CA Previews** file, a "work-in-progress" file containing incomplete references prior to their addition to the CA file. CA Previews records have bibliographic information, author abstracts in many cases, and some CAS registry numbers. The CA Previews file is updated daily and information remains in the file for six weeks after the record is published in the CA file (or the printed CA abstracts).

The **Derwent World Patents Index** (WPI) is a comprehensive patent abstracts file with equivalent family information available on four online systems: Dialog, ORBIT, Questel, and STN. Extensive chemical and polymer indexing is available only by subscription.

The **American Petroleum Institute** (API) Literature, Patent, and Business files are available with full indexing and unlimited usage only to subscribers of products and services produced by the Central Abstracting and Information Services of the API. Searching is carried out most effectively using the powerful API thesaurus. Most patent citations are derived from WPI patent records, so the APIPAT file gives access to Derwent-copyrighted titles and abstracts via API indexing. Literature records are mostly written by API abstracters, but some are taken from Chemical Abstracts or other sources.

U.S. patents back to 1950 are covered with extensive indexing in the **CLAIMS** files from IFI/Plenum. These files are among the oldest online. Text searching is available to anyone, but indexing and reduced online rates are available only to subscribers. Another source of U. S. patent information, but without indexing, is the **USP** file from Derwent, which uses tapes from the U. S. Patent and Trademark Office.

Non-U. S. patent information is available from many specific files. **JAPIO** is an English-language file of Japanese Kokai, unexamined patent applications. **EPAT** and **PCTPAT** cover patent applications filed under the European and "World" Patent Cooperation Treaty (PCT) conventions, respectively. Each is produced by INPI, the French National Industrial Property Institute, and is currently available only on Questel. English-language titles and abstracts are generally included.

The **RAPRA** file is a specialty polymer file with a useful thesaurus and broad coverage of technical and trade literature and conference proceedings. It has a European perspective but also includes good U. S. and worldwide information.

The **Chemical Industry Notes** file from the American Chemical Society and the Predicasts **PROMT** file from Information Access Company each cover the chemical industry business and trade press. Chemical Industry Notes is specific to the industry and contains limited indexing and CAS registry numbers. PROMT has much broader industry and source coverage, which often leads to relevant information being recovered from seemingly "obscure" journals, magazines, newsletters, news releases, or multiclient studies. PROMT also has useful product and industry coding and many sources in full text, including *Chemical Week* and *Chemical Marketing Reporter*.

COMLINE and **PLASNEWS** are each newswire-type files on several online systems. COMLINE is a Japanese newswire that covers several industries, including chemicals. PLASNEWS is produced by *Plastics Technology* magazine and draws its information from three sources: the news staff of *Plastics Technology* magazine itself, COMLINE, and the European Chemical News. COMLINE records are also indexed selectively in the PROMT file.

subject coverage or breadth. Therefore, what we do is offer more customized, personalized services. For this, we develop current-awareness profiles, variously called SDI (Selective Dissemination of Information) profiles or Dialog Alerts, or carry out periodic, searcher-initiated updates. SDI profiles are run against file update tapes or online databases. Initiated updates may be run off-line by database producers, such as the ISS (Individual Search Service) from Chemical Abstracts or at current awareness service centers. SDI profiles may also be run automatically against online databases with special "save search" commands for most file updates on all the major online systems. These profiles are effectively online searches run "off hours," but they retain many of the advantages of other online searches, such as delivery as an electronic, rather than hard-copy, file. However, even these searches do not meet many of the factors of Table I owing to the absence of any searcher intervention. Therefore, searcher-initiated current awareness update searches are run to account for nonstandard update frequencies, structure-based searching, or crossfile searching with duplicate identification and elimination, or when extensive record evaluation or post-processing is required.

Adding value

Some of the various current awareness options may be understood by examination of what is available on the Chemical Abstracts file. The four basic options are summarized in Table 3. At Amoco we take particular advantage of cost-effective CA ISS searches. Most project areas have one or more profiles, and report copies are either ordered in multiples or reports are distributed internally. Yearly profiles are very economical at approximately $700 plus a nominal charge for extra copies. This approach is satisfactory when research areas are well defined, research team members have very similar interests, and hard-copy reports are sufficient.

On the other hand, searcher-initiated, multiple distribution, monthly update reports are very effective for satisfying the needs of larger organizations. Having dozens of recipients makes the unit cost for such value-added searches cost-effective. In one monthly update report using the Chemical Abstracts file, we evaluate approximately 200 CA citations per biweekly file update (more than 400 citations per month) online to display in full about 60 citations of interest to the customers. These citations are categorized to allow readers to focus on particular interest areas. We issue the final report in word-processor format for improved readability. Very reasonable redistribution fees are paid to the Chemical Abstracts Service for all reports sent to more than 50 people. In addition, the retrieved citations are stored in a bibliographic database for archival access. We use Library Master software on our Novell local area network (*6–8*). We have developed similar value-added update reports over the past several years, several of which are described here, along with their relation to the research mission

of the company.

Aromatic Acids Monthly Survey. Amoco Chemical is a world leader in production of aromatic acids, such as terephthalic, isophthalic, and 2,6-naphthalenedicarboxylic acids. Our staff follows technology on oxidation and purification methods. Although we produce no commodity polymers using these acids, we need to understand polymer and other applications, current and potential. The IR&A team publishes a monthly hard-copy and electronic report with broad technical and industry coverage for a whole business unit. Information is taken from the Derwent WPI, Chemical Abstracts, Chemical Industry Notes, and PROMT files, among others. Survey production involves significant reformatting and records categorization and inclusion of researcher comments and contributions. Archival access via networked bibliographic databases is available.

Polyolefins Patent Survey. Our company produces billions of pounds of polypropylene per year at plants in the United States and Europe. Our efforts include catalyst development and licensing, polymer production, processing, and applications development. One current-awareness tool we have developed is a weekly online search of the Derwent World Patents Index file. A Dialog Alert profile is run automatically and the output directed to the searcher's Dialmail inbox. The results are then transferred to a personal computer, where relevant citations are imported into a Library Master database. The records are categorized according to main technological emphasis, such as catalyst, polymer production, polymer composition, or polymer processing. An ASCII text file is output with section headings and transferred via e-mail to more than 50 local recipients, each of whom has access to the archival file upon request.

Chemical Industry Newswatch (CIN). The IR&A team has been producing this weekly product since 1986. The search strategy seeks chemical industry news and general technical issues of interest to company personnel, for whom the CIN is considered to be everything from an information "safety net" to a principal news source.

The CIN started out as a value-added selection of Predicasts-copyrighted abstracts from the PROMT file. It was a simple edited report for about 75 researchers at the Naperville Research Center. As PROMT and personal computer technology has evolved, the CIN has increased its full-text article content, although many articles are excerpted to include broadly relevant information. Records are reformatted and categorized by product or general interest area. Report preparation involves substantial identification and removal of duplicated information because the PROMT file covers many overlapping magazines and news sources.

Distribution of CIN now involves both word-processed hard-copy editions and mainframe documents. Fees are paid to Information Access Company for distribution of the CIN to more than 300 Amoco Chemical Co. subscribers worldwide. The overall cost of the CIN per subscriber has remained stable at about $125–$140 per year. No archival access is available, primarily because of high licensing fees, but information specialist retrieval capability is retained for elimination of duplicate information.

How do you value information services?

As members of the Technology Support Department, IR&A does not benefit from allocated funds or overhead income. Our potential customers do not have any money invested in the department and may or may not choose to work with us as appropriate. Therefore, we compete for limited service dollars with other internal and external providers. This need to be competitive has strengthened our products and services and resulted in loyal customer relationships with the people we serve.

Information is not free, contrary to much common misperception, and our services are not inexpensive either. We are in the same research salary continuum as the customers we serve. We need to recover the costs of our salaries and other operating overhead as well as information retrieval costs, which overall amount to more than $1 million per year for our group. The only costs absorbed by others are database subscription costs for Derwent World Patent Index, IFI/Plenum U.S. Claims, and the American Petroleum Institute technical and business files, which are covered by the Amoco Corporation Patents and Licensing Department and various Amoco subsidiaries, whom we serve.

The value of IR&A services is ultimately determined by the marketplace. We have for years tried to quantify this value and to collect success stories resulting from our work.

Table 3. **Current awareness using Chemical Abstracts**

	Search type			
	CA Selects	Independent search service (ISS)	Online profile[a]	Searcher-initiated
Relative cost	Lowest, fixed cost	Low fixed cost	Medium fixed cost	Higher variable cost
Multiple copy cost	Same as unit cost	Low	Same as unit cost	Low for large distribution
Frequency	Biweekly	Biweekly	Biweekly	User-determined
Searching capability	Catalog topics only	CA file, limited access to profile[b]	All online features restricted to a single answer set	All online features, including multiple or cross-file searching and sample display evaluation
Delivery medium	Paper	Paper	Paper, or electronic file via STNMail, Dialmail, e-mail (e.g., Internet)	All options

[a] Uses Selective Dissemination of Information (SDI) profiles.
[b] Multiple answer sets allowed: no combination of "and" and "link" operators: no structure search.

Unfortunately, these success stories have been difficult to identify and harder yet to quantify. We regularly survey recipients of our newsletters and CA ISS searches. A simple question asked but difficult to answer is "Does this service save you time, and if so, how much?" About 70% of the CIN survey respondents report time savings of well over one hour per week per person on average. This amounts to a corporate payback relative to expenses of between 10:1 and 40:1. Similar results are obtained from the CA ISS surveys.

We clearly see benefit to our customers of all the current-awareness products we offer, and they provide us with regular revenue. We have many loyal customers who come to us regularly for individual information searches. But we do not take our positions for granted. We have developed a marketing presentation highlighting our skills and strengths, including our understanding of our customers' business. Ultimately, our best marketing is the one-on-one rapport we develop with our customers.

The future of information retrieval and analysis

Maintaining currency in technical and market information will involve more information, more media for access, more improvement in accessibility to non-specialists, and more full text but less delay time. There will also be more opportunity to miss important information and more need for trained information analysts. Most information will be available electronically, including journals themselves, and "natural language" query languages will ease access to relevant information for those who wish to search themselves. However, researchers will need more discipline and skills to handle this information. This is reminiscent of the "photocopier syndrome," in which people overwhelmed by print media end up photocopying articles rather than actually reading and analyzing them. Another form of this syndrome is collecting relevant abstracts without actually ordering the original documents. Information specialists can be valuable consultants and analysts for overwhelmed research teams.

One exciting new medium that we have been evaluating involves Lotus Notes Version 3 with the Verity Topic search system. Business and trade information can be sent electronically into Lotus Notes databases by services such as NewsEdge and Individual Company's FIRST!. For example, the entire *Wall Street Journal* is available through each of these services. Similarly, queries to Dialog Information Services files can be made from the personal computer desktop using Lotus Notes via Sandpoint Company's Hoover system. These services all require considerable customer education, another important role for the information specialist.

Improvements in information access and timeliness are being made constantly without extra costs to users. For example, the evolution from abstracts to full text in the PROMT file was not accompanied by significant price increases. In fact, full record display is the same price for PROMT's abstracts and for full text or original article excerpts. Now customers are requesting substantial reduction in the typical 3–5 week delay between article publication and inclusion in PROMT.

Licensing agreements must change

One thing that will have to change is the licensing agreements for information sharing. They will have to be simplified and costs reduced. A model licensing agreement is from Chemical Abstracts, which allows redistribution of citations to up to 50 people in a work area and nearly unlimited electronic storage. We pay very reasonable fees to Chemical Abstracts each January to cover distributed reports of the previous year. We have found other information producers with very high redistribution or licensing fees, in at least one case high enough to prevent us from using their information. Although we information professionals appreciate these issues, most information users will remain oblivious to them and will just share information. Further consumer education and moderation of licensing agreements will be necessary.

As these changes and improvements are made, we will be able to provide our own customers with improved products as well. Rather than being a threat, these new developments will provide us with the tools we need to remain a major source of information at Amoco.

References

(1) Toffler, A. *Future Shock*; Random House: New York, 1970.

(2) Engel, G. *Wescon Conference Record, Vol. 34, 1990*; Western Periodicals: North Hollywood, Calif., p. 535; EI/COMPENDEX *91*(9):113017.

(3) Loftus, F. CHEMTECH, August 1993, p. 12.

(4) Brown, T.; LeMay, H. E. *Chemistry: The Central Science*; 3rd ed., Prentice Hall: Englewood Cliffs, N.J., 1986.

(5) Buntrock, R. E. *Database*, June 1988, p. 111.

(6) Wolff, T. E. *Database*, December 1993, p. 68; *J. Am. Chem. Soc.* **1992**, *114*, 796.

(7) Wolff, T. E. *Database*, April 1992, p. 34.

(8) Wolff, T. E. In *Recent Advances in Chemical Information II*; Collier, H., Ed.; Royal Society of Chemistry: London, 1993; p. 37.

Robert E. Buntrock is a Research Associate on the Amoco Corporation Information Research and Analysis team (Internet e-mail rebuntrock@amoco.com). His B. Chem. degree is from the U. of Minnesota and his Ph.D. (Organic Chemistry) from Princeton. After more than four years as a synthesis chemist (pesticides) with both Air Products and Chemicals and Amoco Oil Co., he joined the Information Division of Standard Oil Co. (Indiana), predecessor to the current group. He has been actively providing current-awareness services for more than 21 years.

Thomas E. Wolff is a Senior Research Scientist on the Amoco Corporation Information Research and Analysis team (Amoco Corporation, Mail Station F-1/CS, P.O. Box 3011, Naperville, IL 60566; 708-420-4662; Internet e-mail tewolff@amoco.com). He received his B. S. in Chemistry from MIT and his Ph. D. in Inorganic Chemistry from Stanford. He spent 10 years in the areas of polyolefin catalysis and oxidation of alkylaromatics to polycarboxylic acids before moving to IR&A. He is a registered U.S. patent agent.

AH, BUT WHICH?

"A good model is one that does not account for all the data, he explains, for some of the data are bound to be wrong."

Francis Crick