(5) Haiduc, I. "The Chemistry of Inorganic Ring Systems"; Wiley-Interscience: London, 1970; Vol. 1, 2.
(6) Armitage, D. A. "Inorganic Rings and Cages"; E. Arnold: London, 1972.
(7) Pantel, S.; Becke-Goehring, M. "Sechs- und achtgliedrige Ringsysteme in der Phosphor-Stickstoffchemie"; Springer-Verlag: West Berlin, 1969.
(8) Allcock, H. R. "Phosphorus-Nitrogen Compounds. Cyclic, Linear and High Polymeric Systems"; Academic Press: New York, 1972.
(9) Heal, H. G. "The Inorganic Heterocyclic Chemistry of Sulfur, Nitrogen and Phosphorus"; Academic Press: London, 1980.
(10) Haiduc, I. "A Survey of the Nomenclature of Inorganic Ring Systems. Proposals for a Unified Nomenclature". *Rev. Inorg. Chem.* **1980**, *2*, 219-49.
(11) Anon. "Changes in Organization, Titles and Coverage of *Chemical Abstracts* Subject Sections for the 11th Collective Indexing Period". *CAS Report* **1981**, *11*, 3-6.
(12) Roesky, H. W.; Rao, M. N. S.; Nakajima, T.; Sheldrick, W. S. "Synthesis of Sulfur-Nitrogen Compounds with Cage Structure". *Chem. Ber.* **1979**, *112*, 3531-7.
(13) Sheldrick, W. S.; Rao, M. N. S.; Roesky, H. W. "Bicyclic Sulfur-Nitrogen Compounds: Molecular Structure of *S,S*-Dimethylpentasulfur Hexanitride". *Inorg. Chem.* **1980**, *19*, 538-43.
(14) Roesky, H. W.; Graf, C.; Rao, M. N. S.; Krebs, B.; Henkel, G. "$S_5N_6(CH_2)_4$ the First Spirocyclic ($1\lambda^6$-Thiacyclopentane) Derivative of a Sulfur-Nitrogen Compound". *Angew. Chem.* **1979**, *91*, 846-7.

# A Pharmaceutical Information Manager's Viewpoint on R & D Information Resource Management†

H. D. BROWN
Merck Sharp & Dohme Research Laboratories, Rahway, New Jersey 07065

Received July 14, 1982

The long time span required for product development and the multidisciplinary nature of research, combined with regulatory constraints, pose unusual problems in information resource management (IRM) in pharmaceutical research and development. Management styles, personnel selection criteria, automated systems, and information content are discussed. Service evaluation is highly subjective and dependent on a satisfied user.

## INTRODUCTION

Two decades ago, new regulatory requirements were put in place by U.S. Food and Drug Administration in the pharmaceutical industry. Much has been written in recent times about "drug lag", the need for patent life extension, and the impact of tight regulations on innovation. It is important to note the types of changes that have occurred because these impact the information management activities. Laubach[1] has quoted figures which are significant. Thus, the increased time for development (2-9 years from 1962 to 1976) and the increased cost ($4-$54 million over the same time) are indicative of the complexities under which both proprietary and published data/information must be collected, evaluated, and disseminated.

Along with the regulatory constraints, other changes have occurred. The hurdles over which new product candidates must pass are now higher because the easy problems have been solved, enhanced safety requirements are in vogue, and new diagnostic/detection capabilities have been discovered. Life is more complex in spite of the enhanced sophistication in methodology.

The development of a new human drug may involve the collaborative efforts of representatives from 30-50 distinct scientific disciplines—bringing a greatly increased scope to the need for data/information requirements.

In summary then, our backdrop involves a multidisciplinary requirement for collecting, indexing, storing, retrieving, evaluating, and disseminating data/information/knowledge over a decade of time between project definition and market introduction. Both proprietary and published information are required from manual and computer-based systems.

## PHILOSOPHY FOR IRM

Different management styles bring different degrees of emphasis, as well as a different set of "buzz" words. If we interpret management as the ability to accomplish results through the efforts of others, then several functions/activities are needed, viz., planning, organizing, leading, and controlling. A brief word about these is in order.

**Planning** needs to be at a number of levels. Most important is the strategic (longer range) type. Here, the environment—technical, organizational and political—needs to be assessed. The mission of the organization, a translation of that mission into perceived information requirements, and a notion of the information-gathering habits of the user population are all important. Some insight into the perceived value of information to management, as well as how (by what means) the transfer should occur, must be determined. For example, one philosophy says, "stock a room with books and journals and let those who wish come to use". Another says a more proactive stance should prevail; i.e., the information ought to be "packaged" to meet individual or small group needs and then be sent to the desks of those requiring it. The latter is obviously more expensive but may also be more effective in dealing with the "flood" of information. Planning under the first philosophy would be quite different than under the latter.

**Organization** may likewise be based on a number of premises. My personal preference is for a mission-oriented or functional organizational scheme. Thus, though many units distinguish between published and proprietary information, the user frequently needs both and should not have to satisfy his/her needs by going to separate units. It is true that security requirements are different for the two types of data, but the utility for problem-solving or decision support may be equal. Widely separated geographic units do require extra site supervision, but failure to recognize the functional similarities at several sites leads to duplication of resources.

**Leading** in an information resource management operation must be concerned with subordinates and with clients. A major problem which all staff managers face is the question of authority. As for clients, we must recognize that we are *long* on responsibility and *short* on authority. Hence, our goal must be to do our homework, espouse useful technology, and be sure

† Presented in part before the Division of Chemical Information Symposium on "Managing Information Services & Systems" at the 183rd National Meeting of the American Chemical Society, Las Vegas, NV, March 31, 1982.

0095-2338/83/1623-0078$01.50/0 

that our "sell" matches their needs. For subordinates, adequate training, concern for a reasonable career path, and guidance with regard to priorities, political traps, etc. must be provided.

**Controlling** in an information resource management setting is not limited to budgetary matters. The "information flow" needs also to be provided for. Policies, procedures, and individual responsibilities must be derived, implemented, and monitored. Under regulatory constraints (e.g., "good laboratory practices"), record-keeping, authentication, record retention schedules, archival protection, and retrievability must be carefully maintained.

The questions of budgeting, resource allocation, and priority setting must be resolved under circumstances which do not permit the assignment of a dollar value to the information packets delivered (since there is no general agreement on the value of information). Some ignore the value question and recover costs (with or without overhead) for purchased items plus personnel time. Others treat the entire function as an overhead to be absorbed by the sponsoring organization.

## TECHNOLOGY ASSESSMENT

Technology assessment in information resource management is an important activity. Not only must one be aware of new techniques, new equipment, and new systems development but one also must carefully choose the time to adopt such developments. Timing is most important in relation to the return on investment to be anticipated. It is crucial to know what technology to choose and when to choose it. Foster's[2] studies related to the S curve in R&D productivity are relevant. The right time to adopt a developing technology is on the rising part of the S curve. It is at this time that further effort will result in the greatest improvement of output, whereas if one waits until the zenith is reached, the chance for return is limited or negative.

## EMPLOYEE SELECTION

Now a word about employee selection in information operations is in order. My preference in training is for a physical science (especially chemistry) first, followed by information science. Some laboratory experience is very helpful in gaining insight into "life at the bench". This also lends an aspect of credibility with scientific users.

Memory—photographic recall as well as associative (correlative)—is a very useful attribute in an information scientist. The best computer is still the human brain! Other mental qualities needed include both analytical skills, for problem dissection, and synthetic skills, for composing the answer in straightforward manner. Accuracy, some degree of candor, and the ability to say "I don't know" all have their place in the mental attributes for our hypothetical information scientist. Those who enjoy learning new facts/new concepts, meeting a wide range of personality types, a never-ending stream of requests, perceiving a problem through its elemental facets, searching diligently through chapter and verse, communicating the results factually but enthusiastically, and ending the day with the anticipation of tomorrow will likely be quite successful as information scientists.

In back of all systems are the *information* resources. Books, journals, reference collections, and proprietary and commercial data bases, along with communication networks/telefacsimile devices to bridge the gap between internal and external stores of information, keep our information resources manager occupied. Usage data (made less painful through automation) is needed to guide selection of what must be on-site as well as to assure compliance with regulatory and copyright requirements. The data/information/knowledge requirements for a pharmaceutical industry laboratory are certainly extensive in both breadth and depth.

As implied above in the section on employee selection, people are one of *the* very valuable information resources. People here means not only the members of the information center but also the scientists. Some bench scientists are willing to share their resources with the center. This can be especially valuable in locating outside contacts for esoteric items. The information center staff needs all the "TLC" normally afforded a valuable resource. Among these are training in new systems, a sense of accomplishment from daily tasks, a psychic reward when a job is well done, and recognition when a major feat is accomplished. As in all service operations, each employee must recognize that the most frequent response from user and management is likely to come following our "goofs". In some operations, the positive recognition is so rare as to lead the framing (at least collecting) the laudatory notes.

## TEAMS

The multidisciplinary nature of the mission, the long time span involved, the crucial requirement for information transfer, the requirement for sharing talent and resources over many projects, and the criticality of timing make teamwork essential. An information scientist is a natural member of the team roster. Both the requirement to process (manage) the data generated by the team and the need for outside information input are readily satisfied by a qualified information resource person.

Some projects (e.g., networking) require that a team of information types be formed. This occasionally leads to problems for the manager because of a tendency of information scientists to desire a lone-wolf operation. "My" system is better or more important, and why should I link it to "their" system. Possessiveness, "turf" protection, and a reluctance to function as part of an integrated network of information nodes must be overcome. This requires selling of the benefits of shared resources, the ability to participate in a larger setting, and the chance for a more significant involvement in management objectives. We all must come to appreciate that our best chance for recognition comes from significant accomplishments and not from protection of our provincial piece of turf.

An important contribution which an information scientist can bring to an interdisciplinary team is perspective/objectivity. If homework is well done, a balanced overview can be perceived by the generalist as opposed to the specialist. By virtue of open information transfer between the two, a more reasonable stance can be developed.

## INNOVATION/VALUE

Innovation is a major goal of R&D operations. Sarett[3] has emphasized the importance of the individual scientist in innovation in R&D. Surely this applies equally well in an information setting. On the R&D side, recent publicity indicates that the U.S.A. has not fared well in comparison to other countries. Thus, the issued U.S. patents dropped 25% in number from 1966 to 1976. Foreign patents increased twofold over the same period. In several major industries (autos, electronics, and antibiotics in the pharmaceuticals) the Japanese outpaced us. A recent article[4] shows why this may be happening. The suggestion is that the sharing of risk, high incentives to a large group of people, careful integration of R&D/marketing interface, group decision making, and a truly cooperative spirit between industry and government produce the desired results. Unfortunately, measurement of the value of our output (data, information, knowledge) is difficult. As with most service operations, the "value received" is in many ways based on subjective measures. Though the easiest measures in our shops have to do with efficiency (larger numbers of units of throughput for the same or less cost) and though we certainly want to do that, our first problem is *effectiveness*. Did we address the right problem? Did the information packet cause any *different* action on the part of the recipient? Was the information even used in subsequent action steps by the recipient? What was the ultimate *value* of the decision to which the information contributed? Was its contribution to the decision substantive or supportive? Was it filed away for possible future use or tossed into the round file?

We first have to do the right thing and *then* learn to do it well. Packets of information (large or small), even delivered on time in neat array, with great eye appeal, have no value unless they are used to solve problems or support productive decisions. Hence, we do need an integrated approach with our users. We need candid feedback regarding the usefulness (*value* if you will) of the data/information/knowledge transferred.

## SUMMARY

An overview of the environment (from my perspective) related to information resource management in pharmaceutical R&D has been presented. Some notions of organizational preference (functional), employee selection (chemist turned information scientist), automation (user friendly, cost effective), and the value of project teams (information transfer) have been noted. Difficulties associated with keeping our innovative tools sharp were observed. Finally, we noted that our bottom line—productivity—should first consider what is useful (effectiveness) and then learn how to do it well—efficiency. Success in the management of information resources depends on the proactive delivery of information packets which find their way into problem solving and decision support for scientists or line managers.

## REFERENCES AND NOTES

(1) Laubach, Gerald D. "Impact of Industrial Innovation in the 1980's—The Pharmaceutical/Health Industry". *Res. Management* Mar **1981**, 9-11.
(2) Foster, Richard N. "Effective R&D Operations in the 1980's—Boosting the Payoff from R&D". *Res. Management* Jan **1982**, 22-3.
(3) Sarett, Lewis H. "Stimulating Technological Innovation—The Innovative Spirit in an Industrial Setting". *Res. Management* Nov **1979**, 15-6.
(4) Gerstenfeld, Arthur; Sumiyoshi, Keyi "The Management of Innovation in Japan—Seven Forces that Make the Difference". *Res. Management* Jan **1980**, 30-4.

# Computer Representation of Generic Chemical Structures by an Extended Block-Cutpoint Tree

TAKASHI NAKAYAMA

Foundation for Advancement of International Science, 3-9-1, Amakubo, Sakura-mura, Niihari-gun, Ibaraki 305, Japan

YUZURU FUJIWARA*

Institute of Information Sciences and Electronics, University of Tsukuba, Sakura-mura, Niihari-gun, Ibaraki 305, Japan

Received September 2, 1982

A representation scheme for generic chemical structures (Markush formulas) is presented. This method is an extension of a BCT (block-cutpoint tree) representation for specific chemical structures, and is called an EBCT (extended BCT) representation of generic chemical structures. A general Markush formula is dissolved into several simple Markush formulas by expanding conditions and nested substitutions. Variable parts of simple Markush formulas are represented in terms of a simple type of substituent group called a generic unit. An inverted file for homologous series and a fragment screening facilitate (sub)structure searches.

## INTRODUCTION

One of the important problems in dealing with chemical structure information by computers is how to implement the representation of chemical structures in computer storage. So far, many representation methods have been proposed and implemented for specific chemical structures.[1] Those methods are categorized to two major approaches: the linear representation and the topological representation. One major reason why so many representation methods have been and will be devised is that we do not have such a complete representation method yet to support various types of substructure searches. We have presented the BCT representation of chemical structures as a method for representing specific chemical structures.[2] The features of this method are that it gives a hierarchical view of chemical structures in terms of blocks (ring assembly) and that those blocks become common structural descriptors for both the topological relation and the fragmentation of each chemical structure. As a result, the BCT representation facilitates substructure searches with flexibility and speed.

On the other hand, there have been few reports about the representation of generic chemical structures, although several retrieval systems, such as IDC[3] and Derwent,[4] are in use for generic or patent information on compounds. In practice, however, it is not always satisfactory to retrieve generic chemical structures by those current systems, with respect to the speed and the flexibility in search or the cost of retrieval.[5]

Recently Lynch et al. have reported a descriptive language (a chemical grammar) for generic chemical structures (in particular, Markush formulas in patents), and it is well suited for the description of Markush formulas, supplanting the description by natural language.[6-8] However, it is not so clear how efficient a search their chemical grammar gives, as compared with other representations of generic chemical structures, partly because the whole chemical grammar and the screening system[9] are still under development. Apparently it is advan-

0095-2338/83/1623-0080$01.50/0 