DOCUMENT RESUME

ED 060 879   LI 003 546

TITLE   Summary of Proceedings of ASIDIC Meeting (Arlington
               Park Towers Hotel, Arlington Heights, Illinois,
               September 26-28, 1971).
INSTITUTION   Association of Scientific Information Dissemination
               Centers.
PUB DATE   71
NOTE   28p.;(0 References)

EDRS PRICE   MF-$0.65 HC-$3.29
DESCRIPTORS   Conference Reports; Copyrights; Data Bases; Financial
               Policy; *Information Centers; *Information
               Dissemination; Information Services; *Information
               Sources; *Information Storage; Interaction; Quality
               Control
IDENTIFIERS   ASIDIC; *Association Scientific Information
               Dissemination

ABSTRACT
               The 1971 semi-annual ASIDIC (Association of
Scientific Information Dissemination Centers) meeting held in Chicago
was designed as a working session in which the topics of most concern
to the member organizations could be discussed in detail. This report
is a summary of those discussions. The broad topic headings for
discussion were: (1) Interactions between data base suppliers and
processing centers; (2) Status reports from each of the participating
centers; (3) Committee reports; and (4) Recommendations to tape
suppliers on pricing structures for data bases, data rights and
copyright statements, quality control, data base backup and
educational and marketing resources. (SJ)

SUMMARY
OF
PROCEEDINGS
OF

ASIDIC MEETING

ARLINGTON PARK TOWERS HOTEL
ARLINGTON HEIGHTS, ILLINOIS

SEPTEMBER 26-28, 1971

ED 060879

SUMMARY

OF

PROCEEDINGS

OF

ASIDIC MEETING

Arlington Park Towers Hotel
Arlington Heights, Illinois

September 26-28, 1971

LI 003 546

PREFACE

The semi-annual ASIDIC meeting, which was held in Chicago, Illinois on September 26-28, 1971, was designed as a working session in which the topics of most concern to the member organizations could be discussed in detail. In order to facilitate the discussions, attendance was limited to full member organizations and to no more than three representatives from each organization. The format of the meeting was a round table discussion based on a list of topics suggested by the Executive Committee. This list was later modified and individual topics were assigned priorities by the attendees.

The initial section of this report constitutes a condensation of the major points which were made during the meeting, and the summary should be read as notes of discussion points, rather than as a presentation or position paper. Many of the topics discussed are so strongly interrelated that pertinent comments were scattered throughout the two-day session, so the remarks for each major topic have been collected together to provide some organization and continuity. There is some redundancy, but, in general, the entire report must be read through to cover all aspects of a discussion topic. The language of the transcribed discussion has been retained wherever reasonable, even though this sometimes leads to discontinuity in the presentation since more than one person frequently contributed to the topic under discussion. No effort has been made to standardize the vocabulary, and term usage does vary, depending on the context of preceding and following discussion. The term "center" should be read generally as any full member of ASIDIC using machine-readable data bases; it should not be assumed to correspond to any particular tape supplier's definition, as these vary in quite significant ways. Similarly, the terms "lease", "license," and "subscription" should be read as "agreement" since these terms, too, vary widely among tape suppliers.

As expected and as desired, differing opinions were frequently presented.
One of the primary purposes of the working meeting was to provide an open forum for
unfettered expressions of opinions. And, the Executive Committee, in editing the
report, has tried to retain these differences. However, it must be recognized that,
as a result of the meeting and the discussions, a unified set of recommendations
has evolved which represents the concensus of opinion of the membership.

SUMMARY OF DISCUSSIONS

ASIDIC MEETING

September 26-28, 1971

Chicago, Illinois

I. Interactions between Data Base Suppliers and Processing Centers

Discussions on a wide range of topics related to the supplier-user interface
for the computer-readable data bases consumed the better part of one day of the
meeting. The topics included such items as provisions of the license/lease/sub-
scription agreements, pricing policies and royalties, quality control and delivery
schedules, constraints on output media and user community, implications of net-
working, and differences in agreements with centers both among data base suppliers
and also by a given data base supplier. Specific points on these and similar
items are summarized below, but it must be recognized that all are interrelated.

A. Pricing of the Data Bases

The topic related to the prices of data bases was one which permeated the
entire meeting. The discussions were not in terms of the dollars involved but
rather the haphazard and inequitable ways in which data bases are presently being
priced.

The tape users are, after all, in the retail handling business, regardless
of whether the use is internal to a company or available to any person who needs
the information. The wholesalers (tape suppliers) need the retailers just as the
retailers need the wholesalers if machine-readable products are to become an
important media for the exchange of information. Therefore, it is to the advantage
of both groups to come up with an arrangement (on pricing) that will be mutually
satisfactory, both to the users and the vendors. At the present time the tape
sales are a drop in the bucket in suppliers' income, but if centers do their job,
five to ten years from now they should be providing a very substantial portion of
the income of suppliers.

It was clearly recognized that the pricing of the data bases should protect the suppliers' investment, but it was also pointed out that the demand for computer-based information services is very strongly price-dependent. Computer-based services can, and in fact are, already pricing themselves out of the available market by establishing prices which are above their assessed value (to the user community). The view was expressed that most suppliers have a misconception that the advantage of machine searching is so significant that the information centers will be forced to buy these tapes, but experience shows that unless the data base is extremely large, manual search is very effective and it may be competitive in cost. This is particularly true if the data base is one which has quality control problems or one which has insufficiently deep indexing (to provide effective computer retrieval). Several instances of data base subscriptions being rejected or cancelled for these and similar reasons were cited.

The present methods of pricing by the various suppliers were repeatedly shown to be haphazard, inconsistent, and in many cases inequitable. Most of the non-government suppliers distinguish between "internal" and "external" use of their data bases, usually charging a flat fee for the former and a fee plus royalties for the latter. However, the definition of "internal" varies greatly as the corporate users are normally allowed to serve all divisions world-wide, including divisions which are incorporated under different names.

During the development of the computer-based services, prices (and other considerations) seem to have been negotiated by each individual center on the basis of the case at hand. One concept that appears to have contributed to the variety in pricing is the concern that use of the supplier's data base by information centers which serve industry results in lost sales of the tapes to those people served by the center. Experience with a couple of proprietary data bases where this point of view was maintained initially by the supplier shows that it is not so. One

company will not buy the data bases themselves unless they can forcast a large
volume of usage to justify the cost of the base. The market of many of the
independent centers is small and medium size industry, which is a market the tape
supplier could not reach himself. With proper pricing there need be no restrictions
on use, since there should be a point determined largely by the volume of usage at
which it would be more economical to become a center rather than to purchase
services. This practice would reflect a special case of a sliding fee scale which
provides cost incentives for volume usage.

The topic of royalties as part of the pricing structure led to a great deal
of discussion with the major problem being the undue administrative burden being
placed on centers by virtue of the wide variety of royalty-type payments made.
Some of these are based on centers' gross sales, and others are based on a per hit
or per reference-retrieved type of royalty. Reporting requirements for the
suppliers do not coincide with the reporting requirements for internal center
operations (or the organizational accounting cycles) with the result that at
least two sets of books are being kept. And, since the units of measure for
assessing royalties vary between suppliers and even for the same supplier for
different agreements and/or data bases (or subsets thereof), it is virtually
impossible (certainly economically unfeasible) to convert the accounting to the
computer. One center reports that the running of the (accounting) software (based
on hits only) for an on-line system runs somewhere between 5% and 10% of total
operating costs. And this center, too, has to handle accounting twice because of
time schedule differences. There was strong feeling that the current practices of
royalty payments were imposing such an administrative and financial burden that,
if continued, the costs should be deductible from the royalties or in some other
way charged back against the supplier. Software development and maintenance

incurred for accounting purposes also has to be taken into account. The reporting, in fact, turns out to be a service to the tape suppliers in that it is a much more accurate record of usage than anything he (the tape supplier) could provide for himself. It gives a very good indication (for merged files) of which data bases are being used, with what frequency and with what utility, and also gives an indication of where the overlapping information is likely to occur. Associated with the administrative costs is the question of frequency of reporting and/or payment of variable charges. These, too, vary from monthly to annually, making the cost of processing the payment requests through financial offices alone exceed the dollar value of the payments.

The use of "hits" as a unit of measure for assessing royalties is not an accurate measure of the usefulness of the information (or lack of it) to the user, especially since the vocabulary of many of the data bases is too general to provide high precision in the answers; it is simply something that can be counted with reasonable accuracy. The application of "hit" royalties to multiple copies of results of standard interest profiles was questioned on the basis that it doesn't make sense to pay a hundred royalties for the same hit which occurs in a standard profile that's merely printed and sent out to 100 different subscribers. And the method of using "hits" as a basis for royalty arrangements is not at all realistic for on-line browsing systems. However, it was pointed out that the cost applied to the user (by the center) includes far more than the acquisitions cost of the data base, and that royalties can be amortised rather than charged directly, unless the royalties constitute such a significant portion of the costs that they have to be passed on to the user.

Both retrospective files and on-line retrieval systems raise additional pricing problems which have not been faced realistically by tape suppliers. Many have no policy for dealing with retrospective collections which are replaced either by new current awareness data bases or by new retrospective collections such as more exhaustively indexed material. Merged, special interest files also represent

retrospective collections for which there are no realistic pricing policies, yet they represent one of the supposedly desired spin-offs of the machine-readable media. Many of these latter files will be used in on-line systems with browsing capability for which royalties based on "hits" should not be applicable. A retainer fee or right to use fee, limited as to the length of time for which it would be applicable (i.e., a lease-purchase plan), was suggested as one method of compensating the supplier. Software pricing structures also may be applicable, at least in part. The typical software rental option, for example, has a rental price which decreases year by year and can be prorated over so many years on a usage basis.

Recommendations which were made during the discussions called for equitable pricing arrangements for all centers, regardless of the type of organization or the affiliation of the individual users served; a consistent "unit of measure" (independent of the price tag applied to that unit) for all suppliers who impose variable fees based on usage; and variable usage reporting and payment on a semi-annual or annual basis.

B. Equitable Treatment of Centers

During the two days of discussions it was quite obvious that there were wide ranges in the pricing and use restrictions which were being applied to the various centers by the individual data base suppliers (exclusive of government suppliers). Most centers reported that they had negotiated modifications to the agreements, some due to technical or administrative considerations, others required by legal departments or purchasing agents. There are indications that some suppliers are negotiating different arrangements abroad than within the United States, both in terms of use restrictions and in terms of exclusive marketing rights. The waiver of royalty requirements for a start-up period for OECD participants was one example.

Use restrictions, especially in terms of output media, also vary from center to center, depending on whether the center serves "internally" or "externally".

It was felt that dual standards between US information centers and centers in other countries (Canada and Europe were specifically mentioned) should be justified by the suppliers, if they do in fact exist, especially in light of legal implications on international dealings of American corporations. It was also felt that there were no differences between a company-based information service and a university or any other type of information center. The function. of both is to provide their clientel with information. Also, the profit or not-for-profit motive of the center should have no bearing on either the price to the center or the use of the tapes.

C. Implications of Networking and Decentralization

The implications of networking, loosely defined as cooperative programs and resource sharing between centers, were discussed throughout the meeting. The picture (with respect to individual centers) is very likely to change in the near future...and the situation will be more the dissemination of information by way of resource sharing rather than a large number of 'stand alone' information centers, each with its own pricing policies. It is expensive to search many individual data bases, and, if a center doesn't have a sufficient number of users for any one data base, it certainly behooves them to trade profiles and to centralize the computer processing. There are certain data bases that are not economical to run any one place because of the small number of profiles. Large retrospective files, too, need to be centralized because of the excessive processing costs and the infrequent nature of retrospective questions. The Integrated Subject File, for example,will be far too big and expensive for handling in many organizations.

There are several examples of experimental "networks" already in operation. The NASA centers have had some experimental experience using both the NASA files and other data bases. One center who runs an internal company service serves all divisions the world over. For most of these divisions, the tapes are sent over there complete for them to run. Actually the central site does the input processing, then teleprocesses the tapes to remote sites for retrospective and current awareness search in that area of the world. Exceptions to the agreements were negotiated with the suppliers where necessary on the basis that it's all part of the same corporation. Another corporate wide information center covers all dependent companies including the manufacturing arm. Processing is presently centralized,, but there are plans to use the company's computer-telecommunications network to ship output over communications lines to other locations and have them do local printing or printing on some peripheral equipment on an RJE type station. The Universities of Pittsburgh and Georgia have an experimental project underway in which Information Scientists at both locations input profiles to a common file for batched searching at the central location (Georgia) with output routed to the appropriate site. Lehigh will be participating in this experiment within the next month or two.

Technical questions which were discussed with respect to networking included the role of the various participants and security provisions. It was felt that the center-user relationship should be retained and that the processing site should be invisible to the user. That is, the computer processing of the service could be done anywhere and not necessarily at another center (i.e., could be a service bureau). The question of profile security was raised. While it was recognized that networking did not afford the same degree of proprietary security as did in-house processing, most of the profiles are not sufficiently proprietary to cause serious security problems. It was also noted that a distinction must be made between accidental disclosure and deliberate disclosure. Certainly there are reasonably adequate provisions that can be made within computer networks where there

are "locking keys" on the terminals and other measures to control accidental disclosure. The deliberate disclosure problem should be left to the courts.

There are several factors associated with "networking" or resource sharing which impact the tape suppliers. It is quite likely that there is going to be a reduction of subscriptions which would also be accompanied by greatly increased usage. Certainly, if networking does come to pass and comes to pass rather quickly, there will be a reduction in the number of individual subscriptions, but the information is going to be disseminated to a far larger audience. Now, some suppliers have proposed additional access charges for those centers which don't buy the tapes themselves but get services from someone else. While one particular center is agreeable in principal to this access charge, they think it should be quite low and would prefer to have it in the form of purchased printed materials from that supplier. The variable fee structure based on usage is particularly attractive in networking to many centers as it would allow them access based on actual usage to which they might not otherwise be able to subscribe (by access fee). It was recognized that pricing on the basis of usage, rather than community served, would likely cause large price increases for some of the corporate centers. One supplier, whose data base was used for an experimental period by the six NASA centers, took the position that all six centers should purchase the tape. They replied that the usage they could see for the next several years was sufficient to cover lease costs only at one center. The matter could not be resolved with the supplier, with the result that all six of the centers stopped machine searching this data base.

Another aspect of center-supplier interface which was raised with respect to networking was the limits of control of the central processing sites on other locations. Can second and third party policing arrangements be enforced? This question is particularly critical when telecommunications links are involved, but

the same problems apply to copyright and distribution restrictions imposed by supplier agreements. One center reported that their legal people insisted on a copyright indemnity clause in their lease agreements.

D. Quality Control and Delivery of Data Bases

The issues associated with the quality of the data bases being distributed ranged across consistency of the data and format, notification of changes, and timeliness of delivery. It was stated that the situation has improved significantly over what it was some two years ago in terms of motivation (on the part of the suppliers). However, retrieval services have now moved from the experimental to a service status, with the result that centers are faced with the problem of supplying a reliable service to their customers. Numerous examples of quality control problems were mentioned during the discussions: many duplications of document records (in a given data base), poor proofing, invalid entries, unannounced changes in control characters, unreadable tapes, wrong recording density, and delayed delivery schedules (as much as three months). The extensive editing which is being done as part of the conversions to search formats should be the responsibility of the tape supplier, not each center. The tape and the printed journals should correspond, which isn't always the case. The items on the tape should be timely and indexing should be sufficiently deep to give the maximum advantage to machine searching and to the specificity that you can get with computer searching as compared to manual searching.

Most of the suppliers currently limit themselves to replacement of a bad tape at no charge. But what they do not cover is the processing costs that have been incurred because of the bad tapes. Replacement tapes are often received too late to be of use to the customer since the information is dated. Delayed delivery also causes problems in the construction or retrospective collections.

The recommendation was made that agreements include a quality control clause which should include a penalty clause if it (the delay) is, in fact, due to the supplier's poor handling. There should be some method of adjusting the cost of the data base to force some compliance to higher quality standards -- some sort of a price concession or consideration for running through bad data that has to be constantly rerun. At a minimum, the lease or license agreement must include a right-to-cancel clause for the center with a prorated portion of the subscription or lease price being returned, if they continually fail to meet quality standards or delivery dates. The data base supplier should also take the responsibility for communicating with his users (centers) when someone finds a problem. Probably the most frequent problem area is the lack of notification of change. Much better experience is reported with those tape suppliers who have some technically knowledge-able person as the direct liaison with the centers, who monitors that organization's changes in terms of how they would affect a center.

E. Editorial, Education, and Marketing Support

Centers are, in essence, the supplier's representative within the company (i.e., user community). They could almost be their suppliers' salesmen the way they have to sell the product, and they feel that part of this task appropriately falls to the supplier. Centers should not have to have a marketing staff or have their personnel spend their time in preparing prospectuses on each data base, stating how it can help them, how they can use it, what's there, etc. Packages of this type could be prepared more accurately by the supplier for his data base than by each individual center. It was mentioned that some materials prepared at centers had not met with supplier approval, but that suitable material was not otherwise available. The lease agreement of one supplier provides for deducting a portion of the cost of any marketing in the center that's done for his data base from the subscription cost of the tapes, which should set a precedent for other suppliers.

One center mentioned that much of their outside use often comes as a referral from the data base suppliers. It would definitely help to have the tape supplier indicate to his subscribership or readership where searching of his file is being done.

Backup documentation in the form of editorial support for the scope and coverage of the data bases is very inadequate. It's not enough just to get a supplier's one page descriptive brochure. The centers who use the tapes for search services require the same documentation to get information out as do the people who put it in. The internal documentation used by suppliers would be appropriate to center operations and should be made available as part of the package with the data bases.

F. Tape Backup

The question of getting backup tapes from another center or getting a tape replacement from another center instead of from the supplier was discussed briefly. The problem seems to be one of delay in getting the replacement from the supplier, often a matter of weeks, when another center might be able to provide it within days. However, in the general situation where all tapes distributed were bad, there is no alternative source to the supplier. Moreover, most centers are converting the tapes to internal formats and few retain the originals beyond a few days or weeks.

It was the consensus of opinion that the responsibility for tape replacements is with the data base supplier and it should remain there in order to make him more sensitive to the problems that have been caused by his distribution of bad tapes. There was also an objection to being charged the full initial costs for replacement of a tape issue that has already been paid for.

G. Data Rights and Copyrights

The issue of data rights and copyrights was recognized as one of great complexity and one of far-ranging implications -- and not likely to be settled by discussions at the meeting. However, there are two specific issues which are of direct concern to ASIDIC members: data rights in the data bases and copyright provisions in agreements.

With respect to data rights in the data bases as reflected in lease and license agreements, the issue becomes one of who owns the information and what it can be used for. There is no disagreement that data base suppliers must be reimbursed for whatever portion of their costs they choose to assign to the production of the tape services, ranging from tape duplication costs to full production costs. However, at this point the analogy with printed materials should hold: royalties are not charged on the use of printed publications, neither should they be charged on the computer-readable version. A library, for years, has had printed material and does not control the number of users that pass through and flip through those pages and receive the benefits of that information. One issue seems to be the ease of copying or duplication - an entire magnetic tape of 50,000 references can be copied by xerox or some other method, it would be a significant economical task. However, one of the major advantages of computer-readable information is its capability for new and varied uses, not the least of which is reformatting and repackaging, perhaps with the addition of new information. Many installations take the raw data or distribution data that comes in and manipulate it slightly in order to have it conform to their accessing techniques or the file structure, and then they add to and enhance it. Is it no longer a problem of copyright because it has been modified? Are tape suppliers able to require return upon cancellation of a subscription? The centers are providing a product and a service not otherwise available, and as long as the tape suppliers' income is protected, there should be no restrictions as to use.

The related issue of copyright was also discussed, both in terms of ownership of the information and in terms of specific requirements in the present agreements. Present agreements which require copyright statements on each page or reference have either been modified by negotiation or are being ignored because they are technically and administratively impractical. One center has taken the position that in order to operate one general software package and stay independent of the data bases, any copyright statement that is required to be on each piece of output, be it a citation or otherwise, must be included in each individual record. They will not print special forms or make the software specific to the data base. Through negotiation, several centers have made arrangements for the citation to stand in lieu of the copyright statement on any output. The question of the legality of the copyright requirement was raised with one participant commenting that it was his understanding that the only thing that is copyright on most of these data bases is the text of the abstract; a citation itself is not copyright. Therefore, the copyright may apply to the data base collection as a whole, but not an individual item.

The group agreed to seek legal counsel on the matter of copyright.

H. Output Media

The restrictions on output media (in present agreements) seem to be involved with the pricing policy and with data rights. The question of pricing for reuse or for third party use is one that worries suppliers. Present lease agreements do restrict us from anything but hard copy, but there are some situations where, for example, tape output would be desirable. Teleprocessing is going on, however, within at least one "internal" corporate center and others are seeking permission to allow transfer of hits through tape media over transmission lines to remote locations. The position was recommended and accepted that there should be freedom

to produce what is considered normal search output in any one of a variety of forms, and that the restriction on copying the data base as a whole be accepted.

I. Agreements

Most of the concerns with respect to present agreements have been raised in the topical sections to which they apply. There are two additional, general comments which are of concern.

The first has to do with the price quotation. The lease or license agreement is the official document governing the agreement between the two parties. Therefore, any additional charges, such as tapes, postage, royalties, etc., must be defined in the agreement. Centers have reported difficulties between tape suppliers and their legal or purchasing departments because invoices do not agree with agreements and/or price quotations.

The second item involves cancellation rights. Legal and purchasing agents of many centers are requiring at least the same cancellation rights as the supplier. Many are insisting on penalty clauses which become operative when the supplier fails to live up to quality standards, delivery schedules, time schedules for notification of changes, and similar points.

II. Status Reports

Each of the participating centers was invited to present brief summaries of major changes or recent development in their centers.

A. University of Georgia

There are two areas you might be interested in that have been major changes in our center in the past year of so. The first has to do with the installations of an experimental try at a network based on computer communications that work fast. We've just completed the software that allows us to support several terminals (up to 15) at any one time on a first-shift basis for input of the profiles into

the direct access file. And we've completed the output handling system -- such that remote centers are supported for dissemination of the output and each center in essence has an RPG capability, in that they can devise their own contents of their output, their format, and so forth. At the present time we have one experimental system installed and one is at the University of Pittsburgh. We've been using the UNIVAC 9200 as the output device. We've been having the input at Georgia with the output going back to Pittsburgh. In addition, we will be working with Lehigh University for input and output. We're trying to get it done through CDC 6400 linkup to the 360. To explain our activities further, the University of Georgia is part of the University System. The University System is composed of 27 state-supported schools, all administered out of one governing board. A few years back, the board consolidated all the computing activity at the University, and we have been working to get both the computing activity and the information activity which we have going to the 27 colleges. We have been working primarily at the schools through the library and through the campus coordinators for the computing activity. A lot of our use within the University system now comes from Georgia Tech and Georgia State, which are located in Atlanta. So we have plans to add one input terminal and multiple output terminals in Atlanta. We are also moving one of the information specialists over there to handle these two schools. Some of the 27 schools are fairly good users and some of the smaller schools only have three or four profiles in the system. We're working closely with the libraries at these schools and are getting to the point where the library will start picking up the tape acquisition costs of our center, we think.

The second area that is a new area of R & D work concerns two large data bases and this is just to make you aware that we are working on them. We now have in-house the Integrated Subject File from Chemical Abstract Service and also the Chemical Abstracts Service Registry file of one and a half million chemical compounds. And we do have work under way on both of these data bases.

IBM Company

Originally we keyed all the abstracts entered into our system. We've been working on tape input for several years, with Engineering Index Compendex and Chemical Abstracts tapes, to make them compatible to our system. We have extended that considerably, adding the NTIS and the AIP tapes. The only keying we're doing locally is IBM Tech Report Abstracts and some of these are keyed in other areas. As far as growth is concerned, (IBM also has experienced a certain amount of austerity), we actually are serving fewer people with specific profiles and more people using the output from each profile, so that total service is expanding. There are fewer retrospective searches because of the costs affecting this group. We charge $14.00 for a search in a data base and 24¢ a hit for SDI hits. Somebody said that's Chinese money, but the controller doesn't feel that way about it.

We've also changed out format somewhat in the current information selection (current awareness) program. The SDI format that started out at ITIRC was the double-card (where you tear off the stub). We're printing now on continuous form paper. We have incorporated automatic addressing so that the address is printed in a position where we can use a window envelope. Then we have a "summary" order card. For example, with Engineering Index we have the entire citation (for each abstract) listed on the summary card(s). This card may be sent to the librarian, appropriately checked to order copies. Another single card gives us the user's evaluation of the abstracts he received. In the several months it's been operating, we have had a number of favorable "unsolicited" comments in terms of the ease of reviewing the abstracts. There's one other feature that adds to the attractiveness of the new format; the paper's much cheaper than the card stock.

Lehigh University

With respect to our operation, originally it was designed as an interactive system to service the eight research centers (now 12) that the University primarily is. We've just gone through a major re-write of the file structure from what was basically a file organized system to a data base management system in which everything is essentially automatic. An input table comes into the system and essentially data are processed automatically from that point, creating all the files, managing all the information, accounting data, and use statistics that are required, etc. We currently run with Compendex, and we found that, with respect to the old system, the structure of the files was wasting too much space. Currently we have a disc capability of somewhere around a billion characters which we are about to exceed, so one of the reasons we revised the structure of the data base was to reduce this. Currently we have about two years of Compendex on-line. Under the old file structure it ran about eight hundred million characters - under the new one it's around two hundred and fifty million characters.

Lawrence Radiation Laboratory

I told you in February in Washington about our chip storage, our fantastic system. That paper has been added to and updated and should appear in The Journal of Chemical Documentation sometime soon. I think I mentioned then that we were considering inverting the file, and we now have enough users to make that worthwhile. We have done that now and are working on searching inverted files and are saving money. It seems to be working out pretty well. We have one of the terminals that ties in with the RECON system and we find that this is a good way to interview scientists. Sit them down in front of the interactive terminal and they can talk to the computer and get answers through it, and we can see what they like and make profiles from it.

We just got a CDC 7600, but it's not connected with our chip storage as yet. We're waiting until the systems people get that organized a little bit better, so we can make changes in our searching system. But we think that will come about in the next few months.

We're running presently with Nuclear Science Abstracts and Chemical Titles. We are also experimenting with purchasing similar services outside. We're buying some from Georgia and Riverside on Chemical Abstracts, and we're attempting to do some very basic studies on the actual costs of running our own data and buying service outside. And, we hope to publish something on this sometime in the near future. We have purchased Compendex and INSPEC also, but they are not running yet. We have developed a generalized file management system for use on the CDC 6600. The name of the system is Master Control, and a paper by Dr. Victor Hampel was presented on it at a meeting two years ago in case you're interested in looking it up. We are also running a couple of internally generated data bases.

University of Pittsburgh

The University of Georgia report mentioned the relationship between Georgia and Pittsburgh, and we are very excited about it and the experiment. The University of Pittsburgh now has retrospective capability, and they've sectionalized their Condensates tapes according to the five major divisions of the Chemical Abstracts Journal, so we can run retrospective searches by division. During the past year our new development at the University of Pittsburgh has been the initiation of the Campus Base Information System Activity being funded by the NSF. The objective is to try and marshall both the academic and non-academic resources available to serve both the academic and non-academic community. We're emphasizing the activity this first year with two demonstration centers right on the campus. One is in the Engineering Library and one at the major University Library. They will include audio tapes to identify the system, its function, and its objectives.

And the in addition, we're trying to initiate an on-line file; I think the

first one we're shooting for is the New York Times file. With it,the user can sit

down and exploit the New York Times file via console. In addition, we are

determined to increase accession to procure a mechanized tape each year for a

number of years. One of the major developments also during the past year was a

decision on the part of the University administration to replace the IBM equipment

we had at the University of Pittsburgh with PDP-10 equipment. That's caused a

great deal of consternation among others with the accompanying shuffle of programs.

III. Committee Reports

  A. Standards Committee Report - M. K. Park

  The data element recommendations and the tape format recommendations were

inherited from Don Peterson's Chairmanship of the Committee. The last thing he did

as part of fulfilling that office was to send out the ballots on both the tape

and data element recommendations, which were to have been returned to me. To the

best of my knowledge, there were 21 ballots sent out to full members, with a

cut-off date of May 15th.

  I received nine responses. On the tape format recommendations, all nine

were "yes" votes, which, in essence, recommended the ANSI Format.

  On the data element recommendations we also received nine responses - seven

of which were "yes", three of those with reservations and comments, and two "no"

votes. So, if we use the balloting conventions that ANSI uses, we would have to

give very heavy weight to the "NOS" and reservations with respect to the "YESES".

In other words, a majority ballot will not pass an ANSI Standard. Interpreted in

this way, we have five dissenting votes and four assenting votes for the data

element standards. The major criticisms of the data elements were that they were

either incomplete or had not been given enough serious consideration in light of

national and international standards that either do exist or are under consideration.

My opinion is that nine votes out of 21 is certainly not a majority in any case,

and even if they had been all "yes", they do not constitute a strong vote for recommending these standards to the suppliers. And, since there were five out of the nine with either qualifications or "nos", I think that definitely sends the data element standards back to the committee for additional work.

The recommendation that I have made, as Committee Chairman, to the Executive Committee is that we redirect the efforts of the Standards Committee, primarily toward trying to influence the national, and perhaps international, standards committees in areas that are the same as the ones in which we are working. In the case of a tape format, we've adopted an existing standard anyway. The Executive Committee accepted this recommendation at the meeting in Kansas City and we have proceeded on this basis. We now have one person working actively on the ANSI/Subcommittee for bibliographic standards, which is precisely the data element list of recommendations. And copies of the draft ASIDIC data element recommendations did go to all members of the Z39 Subcommittee (Z39.4) that's working in this area. In this respect, we are taking an active interest in the question of bibliographic standards. Other Z39 draft standards that have to do with thesaurus rules and technical report standards are being circulated through the Standards Committee members for comment. Secondly, we are in contact now with the national and international standards body UNISIST and have working copies of their documents.

I would suggest that the Committee should continue the activities in support of other national efforts rather than trying to redo in detail the particular set of recommendations that we have now. We will then be in a position to comment item by item on the Z39 recommendations as they come out. And all of the (ASIDIC) Standards Committee members have been put on the review list for the Z39 draft

documents in line with that recommendation. Another thing that was suggested

for the Committee to undertake was to examine each of the data bases which are

represented in ASIDIC, prepare a detailed report of where those data bases

correspond or do not correspond to national standards, and make strong recommen-

dations to each particular supplier for changes in his data base. I think that's

an activity that could be undertaken if people are willing to comment.

Questions and Comments:

The Committee has been depleted with the loss of three members, so perhaps

the Chairman may want to consider adding members.

Membership on the Standards Committee is voluntary and welcomed.
Please notify the Chairman of your interest.

An informal request was received from Tony Kent of EUSIDIC that we consider

recommending their standard for tape format to the membership of ASIDIC.

Peter Shipma was asked to undertake a detailed comparison of the
ASIDIC and EUSIDIC recommendations and report to the Committee in March.

The Chairman prepared a summary of national and international standards

activities related to the ASIDIC interests for circulation to Standards

Committee members. This document, and copies of the proposed recommendations

which were sent out for ballot, are available from the Chairman on request.

The question was raised as to whether or not we could at least publish the

data element recommendations on which there was no disagreement.

The ballot which was distributed called for a "yes" or "no" vote on
the recommendation as a whole.

B. Data Management Committee - Martha Williams

In tabulating the data that was collected in the ASIDIC Survey a little
over a year ago, it was found that the information was so inconsistent that it
is going to be necessary to resurvey by telephone. Once that data is available,
the beginnings of networking can become a little more possible than they are now.
Each of us will have a copy to know what data bases the centers are handling,
and we can contact each other. Those centers who have not done so yet will find
out where resources for sharing may be available.

IV. Recommendations to Tape Suppliers

An ad hoc working group of five participants was assigned the task of review-
ing the current arrangements with tape suppliers concerning royalties, base prices,
retainer fees, and any other matters directly or indirectly related to pricing
structures and to submit a proposal based on preceding discussions for consideration
by the participants. The recommendations were accepted in principal by a large
majority of the attendees, and these points plus other recommendations coming out
of general discussions have been incorporated into the following sections.

1. Pricing Structure for Data Bases

The recommendation was made that the pricing structure for data bases
(perhaps excluding government-supplied tapes) should include the following points:

    (1) Achieve consistency for all subscribers of a particular supplier's
        data base,

    (2) Support the supplier's income base,

    (3) Impose no undue administrative or accounting burdens on either the
        supplier or the center,

    (4) Accomodate current awareness and retrospective searches, both in
        batch and on-line modes,

    (5) Provide a sliding cost scale for increasing volumes of use,

    (6) Achieve consistency between suppliers on the unit of measure for usage,

(7) Enable outright purchase of data bases,

(8) Permit selective repackaging in both printed and machine-readable forms.

On the basis of these considerations, a pricing structure comprised of two parts is proposed: (1) fixed base cost and (2) sliding usage charge.

The _fixed base cost_ would include those production, distribution, and fixed costs that the supplier elects to include in his base cost, including such items as postage, handling, and tapes. This fixed base cost should be the same for all centers except where postage and handling charges differ (in excess) of those included in the base price.

The _sliding usage fee_ would provide for supplemental price charges on fixed package increments by ranges of units, with a unit-price decrease with increasing volume. Centers should be able to buy on the basis of estimated levels of usage with adjustment, if necessary, at the end of the year. The unit of measure which is recommended is a "profile" or search question, the precise definition of which for accounting purposes would have to be defined.

Under the proposed price structure, either or both parts could vary from zero to whatever price the supplier chose to assess based on normal marketing strategies.

2. Data Rights and Copyright Statements

The recommendation is made that the initial subscription charge be set so as to permit outright purchase of the data base. Tape supplier data rights should be protected against duplication of the _entire_ data base content, via any media (including but not limited to machine--readable forms, printed, and microform, and regardless of any alterations of appearance or format of the entire set of material or portions thereof, by legal copyright provisions. Distribution of portions of the data base via any media is unrestricted, providing the tape supplier's

identification (the bibliographic citation) is included and that the distribution

portion does not duplicate the information content of any service, regardless of

media, normally produced by the tape supplier.

3. Quality Control

It is recommended that clauses be added to the legal agreements between

centers and tape suppliers which would, as a minimum, provide cancellation rights

with the return of prorated acquisition costs of the data base when tape failures

or delayed deliveries are attributable to errors on the part of the tape suppliers.

It is further recommended that a three month lead time be required for notification

of changes in the tape or data specifications, including printed or machine-readable

examples as appropriate, the lack of which notification would be subject to the

same cancellation rights as above.

4. Data Base Backup

Each tape supplier should maintain a complete tape archive of his data

base(s) and provide replacement service at or near duplication and handling costs.

It is also recommended that centers be able to obtain replacement tapes (but not

new collections) from other centers provided the tape supplier is notified of the

replacement. This policy of replacement would apply only to exchange between

centers who have paid the initial purchase price to the supplier for the specific

tape(s) being replaced.

5. Educational and Marketing Resources

The tape suppliers are requested to prepare and to make available

educational and promotional materials which accurately describe the information

content of their data bases. Such resources, which could be made available for

purchase and/or rental, would include visual and audio-visual materials, as well

as printed matter.