

ECONOMICS OF DATA BASE PRODUCTION, SEARCHING EXPLORED DURING SPRING '78 MEETING IN NEW ORLEANS - INTEREST HIGH AMONG ATTENDEES TO COPE WITH TRENDS, CHANGES OF UPCOMING YEARS

The Spring 1978 ASIDIC meeting held in New Orleans April 23-25 has been hailed by many attendees as the most lively and significant meeting of the Association in memory. Even with the new "two-full-days" format, nearly everyone stayed for the entire program that explored all aspects of the meeting theme, "Economics of Information Transfer."

As a change in procedure, the ASIDIC business meeting, open to members only, was held first. ASIDIC is financially viable and has several committees carrying out important projects. These will be reported in the story about this part of the program.

On Monday morning, the Economics of Primary Journal Publication and the Economics of Batch Searching were presented by Robert H. Marks and Dr. Daniel U. Wilde. Marks reviewed the elements of cost in primary journal publication. He divided these costs into two principle areas: 1. editorial and composition; and 2. printing and distribution. Dr. Wilde gave three scenarios of the economics of batch data base searching: 1. Using out of house services; 2. Utilizing a variable cost computer system; and 3. Owning a computer with a fixed cost base of operations.

After lunch, Collin K. Mick presented the Economics of Online Data Bases. Herbert A. Nobles talked about how IBM's ITIRC operation prices its services to users within IBM. The final presentation of this group was given by Donald T. Hawkins of Bell Labs. Mr. Hawkins reviewed the cost effectiveness of manual and online searching and reported on the operations of Bell Labs approaches.

By the time the discussion groups convened under the direction of Doug Ferguson, Kay Durkin, and Greg Payne, many of the attendees expressed the desire to spend some time "digesting" and reviewing many of the things that had been presented prior to launching into evaluative discussions. Nonetheless, all discussion groups proceeded with lively and revealing exchanges of thoughts, ideas, and questions about the presentations of the day and about aspects of the general topic of the conference.

(Continued from Page 1)

The second morning was highlighted by Ronald Segal of EDUCOM as he talked about the Economics of Networking. Mr. Segal explored the technical aspects of networking as a concept and as some functional systems carry it out today. He pointed to the future technologically with optimism, then interspersed the political and the economic realities of developing viable effective networks to show that the "problems" lie more in deciding who will pay for networks and then administer them than in whether or not "they" are technically viable to develop.

Tuesday afternoon saw the previous presentations, discussion points and questions, and many previously closely guarded Areas of the economics of data base and information publishing and distribution surface in a lively panel on the "Economics of Data Base Production," consisting of Tom McDonald, John Creps, J. Ron Smith, ASIDIC President, and Oliver Ball. Following presentations by each of these well-known representatives of important firms

engaged in producing data bases, the entire meeting was opened up for broad ranging questions, answers, rebuttals, and what one attendee described as "the greatest official off-the-record exchange of meaningful ideas this association has ever seen."

Many of the points covered in the open discussion cannot be relayed to those ASIDIC members and friends who did not attend, partly because of lack of space and partly because no one could take copious enough notes to adequately report on the meeting. It is interesting to note that associates in the industry were still discussing the meeting weeks afterward during lulls in the ASIS Mid-Year meeting and the SLA annual meeting.

The only disappointing note to the meeting was the news that Peter Schipma, host of the Fall '78 meeting in Chicago, had to undergo back surgery and could not be in New Orleans to allow the membership to select the final dates and the hotel for the Fall meeting. (Peter is reported doing better and will tend bar at the hospitality mixer in Chicago.)

PAGE CHARGES JUSTIFIED, NECESSARY FOR PRIMARY PUBLICATION SURVIVAL IN STI FIELD, SAYS AIP'S MARKS

Primary journal publication costs fall into two divisions: 1. Editorial and composition, and 2. Printing and distribution, according to Robert H. Marks of the American Institute of Physics.

Before explaining these two cost elements, Marks carefully went through the revenue sources to the publishing effort. Of special interest was the detailed history he gave the practice of "page charges" or "page fees," the practice of soliciting money from a personal author or institutional sponsor of the author's work for each page of an article to be published. Page charges have been a regular practice of the publishers of STI materials since the 1920's, according to Marks. He said that about 85% of all authors honor the page charge plan. Those authors that do pay get their

| CONTENTS | |
|---|----------------|
| ISSUE No. 32 | |
| | <u>Page(s)</u> |
| Spring '78 Meeting Report | 1 - 2 |
| Spring Meeting Speakers | 2 - 7 |
| Fall '78 Meeting in Chicago | 7 |
| Spring '79, Fall '79 Meeting Sites, Dates Confirmed | 7 |

(Continued)

(Continued from Page 2)

articles published more quickly than those who choose not to pay the fee. The delay for nonpayers is up to about three months and does not represent outright or conscious discrimination against this group of authors.

The underpinnings of page charges are economic. "In the 1920's most STI publishers faced extinction if they didn't come up with alternative income sources to supplement markedly lower subscription fee levels caused by the general depression in the United States and the Western World," said Marks. There were three moral factors involved: 1. The completion of the research cycle by publishing and disseminating the findings of projects and experiments; 2. The service that the publishers provided to authors and their institutions by carrying out the publishing and dissemination function; and 3. The instructional service furthering education in the field. Marks said that the reasons for page charges remain essentially the same today.

In 1975, the AIP produced 24,346 pages of editorial material at a cost of \$1,670,300. The cost per page is the result of dividing total costs by total pages. This gives a \$68.60 cost per page.

Marks reported that the costs of getting a publication "to pages" (ready to print) are essentially fixed costs. Composition and printing are roughly half of all publishing costs. These are the areas where new mechanical and computerized methods represent potential cost savings which may be realized.

WILDE CITES THREE ALTERNATIVE DATA BASE UTILIZATION MODES: ECONOMICS ANALYSIS KEY TO BEST CHOICE

Dr. Daniel U. Wilde, of the New England Research Applications Center (NERAC), said that the relationship between prices and costs in the batch searching of data bases is difficult to firmly establish.

Setting into perspective several terms used widely in the industry today, Dr. Wilde

suggested that it is most useful to think of sequential data base searching performed by batch systems while inverted files created from sequentially (usually) ordered data base tapes are most widely used for real time interactive searching systems applications. The end result of good information retrieval is shared by both approaches.

Wilde pointed out that there are three options available to organization that want to benefit from bibliographic data base searching: 1. Buy the services provided entirely by outside firms; 2. Lease a tape of the data base(s) desired and pay a third party to process or "put it up" for you; or 3. Lease the tape(s) and process on your own computer - owned or leased.

Using graphs and economic charts, Dr. Wilde demonstrated the cost, price and volume relationships between each of the options. He then showed the results of the costs of all the options together.

FULL TEXT TREND A SURE THING SAYS COLLIN MICK: DATA CAPTURE TO BE GREATEST CHALLENGE NEXT FEW YEARS

Economics of Online Data Bases was the theme of the presentation made by Collin Mick of Applied Communications Research.

Over the next fifteen years Mick predicts that there will be a trend toward a closer tie-in and possible transition toward full text bibliographic data bases rather than the current abstracting and indexing files so prevalent today. Labor problems and scarcity of skilled data base people, coupled with fast rising costs of all forms of labor, will be very prominent in the industry. The economic issues of cost recovery of the classic "first copy" and copyright protection will be very important to publishers. Close communication and understanding among data base developers, publishers, and online vendors will be needed to avoid serious complications and problems.

(Continued)

Mick feels that the trend toward some reduced costs of online searching may be temporary and may soon turn toward higher levels, resulting in increased prices to end users. Increasingly similar interfaces for search systems and locally available storage and editing capabilities under end user control are realities today. The advent of micro-computers with small diskette (the so-called "floppies") and cassette storage media make end user search and editing control more possible than ever.

The greatest area of unsettled economics in data base production in the next few years will be found in the costs and methods of data gathering or data capture, getting material into machine readable and processable forms to go into the batch or online searching systems.

IBM 'MAKES MONEY' FROM INFORMATION SERVICES IT BUYS ON OUTSIDE: ITIRC DIRECTOR NOBLES TELLS, SHOWS HOW

IBM has carefully measured the payback from the information services that it acquires from outside suppliers or produces internally and makes available to users of ITIRC, the IBM Technical Information Retrieval Center. As measured in hours of labor saved (translated to dollar equivalents) by ITIRC customers themselves, IBM is realizing a five-to-one payback on every dollar spent on information services.

Nobles explained that ITIRC's services are not free to its users. "ITIRC is a cost recovery department," said Nobles, and ITIRC "must keep all costs allocated." Each year, he explained, a budget for ITIRC expenses is set to reflect the cost of acquiring and processing all of the information services it uses. To these costs are added the costs of administering ITIRC. The sum of these costs is the major input into the ITIRC services pricing structure. ITIRC then goes to the various divisions to "sell" them on the ITIRC services. IBM allows its divisions to go outside if they so desire. This puts ITIRC into

a quasi real company situation when it comes to selling and servicing its users, since there is no company policy requiring divisions to buy ITIRC services.

Even though IBM users don't remit "hard" dollars to ITIRC, costs for information services are budget line items that go through regular review and are sensitive to the same considerations as all outside services the company seeks.

Nobles used detailed overhead transparencies to show how the budgets are developed for each ITIRC service and are then input into the internal pricing systems. Any interested members can contact Mr. Nobles at ITIRC for further details.

MACHINE VS MANUAL SEARCHING NOT TOPIC NOW SAYS HAWKINS - BELL LABS EXPERIENCE FIRM WITH ONLINE MODE

Donald Hawkins of Bell Labs in Murray Hill, New Jersey, said that one of the greatest early differences between manual and machine bibliographic data base searching was that in machine searching (online) the firm got "hard bills, ones that had to be paid."

Hawkins spent more time discussing how Bell Labs accounts for computer-based searching than in comparing manual and machine searching costs. As he stated to the group, there really isn't much to compare - online searching has proven itself to the point that the major effort is spent accounting for it, not defending it.

Several accounting techniques, taking into consideration tangible and intangible elements, were presented. The methods used by Bell Labs were reviewed in a series of overhead transparencies.

One important measure of online cost effectiveness can be derived by comparing the cost per relevant document or citation retrieved with the cost of overall retrieval.

(Continued)

(Continued from Page 4)

This figure is derived by subtracting the cost per retrieved item from the cost per relevant item retrieved.

In 1972 Bell Labs accessed two online data bases, did about 25% of their searching via computers, did none of their retrospective searches by machine only, had an in-house searching system, and executed 100 searches by three professional information personnel. In 1978 Bell Labs actively accesses over one hundred data bases; nearly all searches involve online; most searches are "machine only" in nature; they no longer have an in-house search system; and, in 1977, three searchers did over 1,100 searches. The end result, says Hawkins, is increased productivity of information personnel because of online searching capabilities.

SEGAL REVEALS EDUCOM PROGRESS IN DESIGN, MODEL TESTING FOR EDUNET ACTIVITIES: BIG QUESTION - WHO PAYS?

Policy issues, not technological problems, are the main barriers to the introduction and utilization of extensive resource sharing in online or batch networks for educational institutions, according to Ronald Segal of EDUCOM.

Economics, the lack of ready funds to pay for services in "hard" dollars, plagues most schools. There seems to be plenty of computer power to be shared but figuring out who pays and in what ways is a major problem that hinders further development of EDUNET.

Segal explained the technological aspects and trends in the networking areas. Some decreases in communications charges, with the possibility of distance independent costs, are foreseen. Computing power--cpu and storage costs--appears to be declining on a transaction basis but still represents significant shares of the cost of networking. Distributed systems and shared processing seem to work in some cases but have not been proven on a large scale throughout the nation.

The libraries and the data processing centers of schools must determine who and how the costs will be allocated and borne budget-wise in order for real progress in networking to occur.

The simulation project carried out with the National Bureau of Economic Research was reviewed. In this program, actual managers, policy level and operations managers, were brought into the gaming exercise which has proven valuable in gaining insights into the behavior of managers using the proposed or simulated network.

PANEL ON DATA BASE ECONOMICS SPURS MANY QUESTIONS, COMMENTS, IDEAS DURING MEETING WIND-UP SESSION

The last formal part of the Spring '78 ASIDIC meeting was a panel on The Economics of Data Base Production, featuring four well-known executives from the industry: Oliver Ball of INSPEC, John Creps of Engineering Index, Tom McDonald of Chemical Abstracts, and J. Ron Smith of BIOSIS. Each panelist gave a short presentation which was followed by discussion from other panelists and questions from the attendees.

This panel bridged naturally into the wind-up portion of the meeting, the lively discussion covering the entire scope of the different subjects touched upon throughout the two days.

BALL CITES SUBSIDIES

Oliver Ball, the marketing director of INSPEC, stated that nearly all online bibliographic data bases are subsidized either by direct grants or investment or because the online version of the data base is viewed historically as a by-product of a predating print A & I service.

Ball said that rough approximations of INSPEC costs could place a value for each record added to the data base at about

(Continued)

(Continued from Page 5)

ten dollars (American). Since INSPEC inputs about 150,000 records yearly, this represents a cost of about \$1,500,000 to produce the data base. Taking away all of the print product revenue from INSPEC, at an online royalty of \$20 per hour (not a real figure for INSPEC), it would require 75,000 connect hours annually to make the operation viable. Currently, INSPEC receives about 4% of its revenues from online sources. Tape leases account for another 19%, leaving print products generating 77% of the revenue for INSPEC.

Mr. Ball said that online revenue generation and print product sales are not the same. "We used to know that we could 'buy' x number of print subscribers in a given year with a certain marketing and sales campaign. Online is not sold that way. Online hours are sold through education and user service. The concept of renewals in online marketing is hazy at best."

McDONALD CITES CA's STRENGTHS, WEAKNESSES

Chem Abstracts is very good at producing data bases and derivative print products utilizing some of the most advanced computerized techniques possible, according to Tom McDonald, Sales Manager, CA. Chem Abstracts' major weakness lies in the way it carries out marketing programs, not direct sales. McDonald clarified some of the differences between marketing and sales, citing research and product design based upon end user studies and understanding as movements that are starting to happen at CA. "As with most firms, once a weakness is recognized and appreciated, moves to change come quite rapidly. We are now seeing this at CA. Of course, it will take some time."

Attention of management level executives to the costs and benefits of scientific and technical information services in their organizations is increasing faster than in the educational sector. Libraries are becoming viewed as technical information centers that have regular and important input into the process of development in the firms. The need to focus attention on the management of the heretofore fragmented collection and utilization of information is being rec-

ognized, and Chem Abstracts plans to pay attention to the research and results of studies in this area.

Turning an organization from a sales orientation to a marketing concept, where the products and services are derived from end user need research, is not easily or quickly accomplished, but it must be done by all firms engaged in the data base industry, according to McDonald.

CREPS CONFIRMS BALL'S CONTENTIONS

John Creps, head of Engineering Index, supported Oliver Ball in the proportion of expenses and revenues that EI experiences relative to data base development and online searching income. To just prepare the data base costs EI about \$945,000 annually. It costs about \$600,000 in administrative, marketing and accounting expenses that EI would incur whether or not it engaged in online or published information services.

Creps cites a shift in the market demand away from printed books and the A & I journals produced by EI. This makes online become a problem from a management and an accounting standpoint. The challenge is how to allow for online usage growth while shifting the revenue sources and operations of EI from a predominantly print orientation to a predominantly data base orientation. The impact on higher online charges and a realignment of staff and accounting methods is obviously a situation that will take careful and creative management.

DISCUSSION IDEAS FLOW FREELY

Following a short break, the open discussion period that ended the meeting took place. There were many animated exchanges of ideas, clarifications of previously presented points, and a general review of the entire meeting.

Questions such as "Does this mean that \$200 connect hour fees are possible?" were openly discussed. Based upon the economics discussed earlier, it was thought possible that connect time rates could escalate if and when print revenues began to decline

(Continued)

markedly for some of the large data base suppliers. Very few data bases were founded as computerized products where the impact of on-line growth will be wholly positive.

It was suggested that the rapidly growing number of online user groups might find a collective home and forum in ASIDIC more to their liking than the Special Libraries Association (SLA) or the American Society for Information Science (ASIS). It was suggested that some of these groups might be interested in learning more about ASIDIC. The Houston users group sent a representative, Mr. A. O. Zike, to New Orleans to investigate such possibilities. Aspects of where the best forums for meeting and planning with data base systems and producers were explored and ASIDIC was suggested as a prime location for such meetings.

Attendees left New Orleans with the impression that the industry faced several challenges in the next few years that will make ASIDIC even more important as a forum for meetings on the changing nature of the data base development and publishing business. The online systems and the tape processing centers closely coupled with the rapidly growing number of data base producers.

FALL '78 MEETING CENTERS ON NEW STANDARDS: DB DEVELOPERS, SUPPLIERS, SYSTEMS USERS ALL FEEL IMPACTS

The Fall ASIDIC Meeting will be held October 29, 30 & 31, 1978, at the Blackstone Hotel in Chicago, Illinois. IITRI, under the direction of a much healthier (following back surgery) Peter Schipma, will host the meeting.

The meeting will center on the impact of emerging international standards upon the production, systems handling, and end user utilization of bibliographic data bases. "International Standards are no longer merely a topic of dull papers and a matter of routine correspondence from the ASIDIC Standards Committee," commented Dennis B. Auld, Vice Presi-

dent of Operations for Data Courier, Inc. "Data Courier has spent many thousands of dollars preparing to operate within the guidelines of these standards. These expenditures have impacted our editorial, production, and marketing efforts greatly, and we expect our adoption of the standards to pay off for us over the long run."

The final program theme and schedule will be sent out in the ASIDIC Newsletter in late August. Margaret Park, head of the ASIDIC Standards Committee, is working closely with the Program Committee and the ASIDIC Executive Committee to formulate a program that will elicit participation and meaningful results (adoption of an ASIDIC position on the amended UNICIST Reference Manual) for all attendees and the association.

MARK YOUR CALENDARS NOW FOR SPRING AND FALL 1979 ASIDIC MEETINGS: OTTAWA, BOSTON SITES SELECTED AND CONFIRMED

The Spring 1979 ASIDIC Meeting will be held in Ottawa, Canada, May 26-28. Hotel and meeting theme details will be available in the next ASIDIC Newsletter.

The Fall 1979 Meeting is slated for the Parker House in Boston, September 16-19. NERAC, headed by Dr. Daniel U. Wilde, will be the host organization for that meeting.