Grant Project Information via a Shared Data Base

Justine ROBERTS: The Library, University of California, San Francisco

A quarterly keyword index to campus grant projects is provided by the Health Science Library at the University of California, San Francisco, using a data base created and maintained by the campus' Contracts & Grants Office. The index is printed in KWOC format, using the chief investigator's name as the key to a section of project summaries. A third section is also included, listing the summaries under the name of the sponsoring department.

INTRODUCTION

Communication channels between the computer center and the library at the University of California, San Francisco are open despite the "normal" and accompanying library use traumas of an all-purpose university computing center. Thus the library's chief administrator received an immediate, if unexpected, response to her statement of campus need for subject access to information about local research and training projects. As she summarized it, the information need is expressed as "Who is doing what, where, with what amount of funds?" Such gueries about campus work come to the Health Science Library regularly, but had often remained unanswered because of inadequate published sources and the lack of easily accessible local sources. Neither the campus Contracts and Grants (C&G) Office, nor any other campus unit, had files organized to allow inquiry by subject or by department name, nor were any departments staffed to provide a general information service of this nature. Previous investigation by the library had revealed the fiscal infeasibility of extracting citations of publications on campus research projects from a commercially available data base. The latest locally compiled directory of campus work was eleven years old.

Response by the computer center director to the library statement was a suggestion for a three-department cooperative project between the library, the computer center, and the C&G Office to produce a quarterly index of the machine-readable administrative file of the latter department. This accession-ordered file is comprised of 1441-character records which provide for 42 data elements needed by the C&G Office to monitor the progress and fiscal status of all extramurally funded projects and proposals (Figure 1).

052771070167063072070170063072	``	TO1 AH05	1G404961P1FDRSHAM	0-1000940 004	
00000 00000000 00000000 00000000 000000	00 000000		DIABETES	IN ENDOCRINOLOGY	RAINING GRANI
00224146 1430005496100050890000000000005221300048346	2000000002	00000000 00000	00000000 00000000	00000000 000000	00000 000000
000000000000000000000000000000000000000	00 0000000	574000000000	0000000000049402000	5300054944000508	000000002414
000000000000000000000000000000000000000	0000000000	00 000000000		000000000000000000000000000000000000000	000000000000000000000000000000000000000
000000000000000000000000000000000000000	0000000000	000000000000000000000000000000000000000	00 0000000000	000000000000000000000000000000000000000	000000000000000000000000000000000000000
00 000000000000000000000000000000000000	0000 0	000000000000000000000000000000000000000	0000000000000000000	00 000000	000000000000000000000000000000000000000
7070172063073 0000000000112742050010439116001069361	7206307707	00000014072871	0000000000000000000	000000000000000000000000000000000000000	00 0000

Fig. 1. Dump of Contracts & Grants Office Master Tape Record

Original specifications for these project records had in fact included a gesture toward information retrieval in the form of a 5-digit "discipline" code; this code had quickly become null when problems of maintenance and interpretation revealed themselves. However, the regular monthly entry and updating of other descriptive data was already twelve months underway at the time the cooperative project was suggested.

LIBRARY INDEX

The original proposal for a library index to the C&G file was production of a standard KWIC (keyword-in-context) index to project titles, to be based on use of an IBM SHARE library program developed by computer center staff. After review of available library programs and output, the product was finally specified to be a KWOC (keyword-out-of-context) index to project titles, using the chief investigator's name as key to a second "bibliographic" section of project summaries (Figures 2, 3). A third section was added to list project summaries indexed by campus department name (Figure 4).

The C&G file included 12 elements which the library considered to be of general campus interest. These elements, comprising the project summaries, are: (1) project title; (2) chief investigator's name; (3) award status (i.e., funded or proposed); (4) project site (i.e., campus or affiliated institution); (5) project type (e.g., training, basic research, applied research); (6) grant number; (7) total project duration to date; (8) award period; (9) award amount; (10) granting agency name; (11) campus department; and (12) school.

Items 10, 11, and 12 of this list exist as numeric codes in the records,

UCSF LIBRARY	UCSF CUNIKACIS & GRANTS INVEX . 9/14/73	PAGE 29
MENTAL		DUNLAP. N SCHLESINGER. H
MENTAL	MENTAL FEALTH SERVICES TO LEAF REMABILITATION PATIENTS	SCHLESINGER, H
METABOL	SUBCELLULAR MECHANISMS OF PLASMA LIDOPROTEIN METABOL	HAMILTON, R
METABELIC		SINGER, T FISHMAN, R
METABOLIC		DCKNER, R
METABOLISM	STUDIES OF BILE PIGMENT METABOLISM	SCHMID, R
METABOLISM	BILIRUBIA METABOLISM DURING PERIATAL PERIOD	THALER. M
METABOLISM	METABOLISM OF BASIC AMINES OF THERAPEUTIC INTEREST	MAY, E SMITH, L
METABOLISM		WATSON. J
METABOLISM	REGULATION OF LIPID METABOLISM IN CULTURED MAMMALIAN CELLS METABOLISM OF PSYCHOTOMIMETIC-PHENYLISOPROPYLAMINES	CASTAGNOL 1 · N

Fig. 2. Keyword Subject Section

UCSF LIBRARY	. UCSE CONTRACTS & GRADIS INDEX 9/14/73	PAGE 79
RCSENAU. W	LYPERCCYTE & TPMUNE CYTCLYSTS PRCJ TYPE B SITE:UCSF AWARD PATHOLUGY PR 11 L 9/C1/66 TO 12/31/74) AGENCY GC21 NO: RO1 CAO719	(CONTINUED)
ROSENAU. W	TRANSFER FACTOR-CHARACTER & ROLE TUMOR IMMUNITY PATHOLOGY PROJ TYPE B SITE:UCSF AWARD	
POSENTHAL. F	YR 1 (7/C1/73 TO 6/3C/74) \$12,002 AGENCY 5C36 NO: 13 EXTRACELLULAR POTENTIAL FIELDS OF CONTICAL PT-NEURCHS	
	CARDICVASCULAR RES PROJ TYPE B SITE:UCSF AWARD AGENCY 0061 NO: ROINS-0974	6
PCTHMAN, S	SECRETION PROTEINS & ICNS BY PANCREAS DENTISTRY PROJ TYPE B SITE: UCSF PROPOSAL	2 miles (1 miles)
RUDNICK, A	THE FOREST CYCLES OF DENGUE IN MALAYSIA MCCPEP FEUDICATION PROJ TYPE B SITE: UCSF AMARD YP 2 (6/01/73 TO 5/31/76) AGENCY 0036 NO: R22AI11540	

Fig. 3. Chief Investigator Section

with decoding tables comprising a separate file at the beginning of the tape volume. These items, together with items 3, 4, and 5 are coded on input but are not uniformly edited by the regular C&G update program.

PROGRAM REQUIREMENTS

Necessary program functions included the selection of active grant records and the editing, decoding, and formatting of selected data elements for printing, and the extraction and sorting of index terms. These functions were divided between a main routine coded by library staff, an indexing subroutine and print program written by a computer center staff member, and an IBM utility sort. Local programs were written in PL/1, using the newly installed PL/1 Optimizing Compiler, and provided several tests of the compiler's capacities.

Only projects currently in "award" status, or which have outstanding proposals during the preceding twenty months, are selected for printing, approximately two-thirds of the file at this time. These conditions are tested on various data fields in the C&G record, and data from selected records are reformatted into a partial print line and passed to the keyword extraction

UCSF LIHRARY	USSE CONTRACTS & GRANTS INDEX9/14/73	PAGE 104
MEDICINE, GENERAL .	INTESTINAL ARSCHPTION OF FEMILECENE 1PEN SCHOOL OF PENICINE PROJ TYPE H SITE: UCSF PROPOSAL	(CONTINUED)
MECICINE. GENERAL .	REGULATION OF RICTRANSFORMATION IN THE LIVER SCHOOL OF MEDICINE PROJ TYPE H 'SITE: UCSF PROPOSAL .	BISSELL, O
MECICINE. GENERAL .	PATHUGENESIS OF ACUTE HENCE FAILURE SCHOOL OF MEDICINE PROPOSAL PROPOSAL	DAUGHARTY, T
MEDICINE, GENERAL .	HUMAN THYROID CARCINCGENS SCHEDL OF PEDICINE PROJ TYPE B SITE: UCSF PROPOSAL	OKERLAND, M
PEDICINE, GENERAL .	LYSYL CXIDASE RECIPTED CYLLACEN CYPSS LINXING SCHOOL OF MEDICINE PROJ TYPE 6 SITE: UCSF PROPOSAL	SIEGEL, R
MEDICINE, GENERAL .	CONTRACT TO DEVELOP & CONDUCT TRAINING SESSIONS FOR PERO PERSONNEL SCHOOL OF MEDICINE PROJ TYPE T SITE: UCSF PROPOSAL	MATTS. M
MEDICINE, GENERAL .	PEGULATION OF GFAL EXPRESSION BY PECFFIRE STERGID COMPLEXES SCHOOL OF MEDICINE PROJUCT TYPE B SITE: BUSSF PROPOSAL	BAXTER, J
METABELIC UNIT	STUDIES IN IASULIN RELFASE IN GLADETES SCHOOL OF REGICINE YR 5 (1/01/07 TO 6/30/75) 46,434 AGENCY 0041 NOT ROLAM12763	FORSHAM, P

Fig. 4. Departmental Index Section

routine. This section checks each word of the project title against a stop list and duplicates elements to provide fully formatted index print lines for each Index section. After sorting, the Index sections proper are printed by the separate print program. Front-of-Book Code and Abbreviations lists, error reports (Figure 5) and a stopword list, are printed during the main routine. A copy of the error report section on codes is sent to the C&G Office following each run, together with a complete copy of the Index.

The need for standardization of decoded values in the C&G tables proved to be minimal, primarily involving limited inversion of department names, e.g., General Medicine is changed to Medicine, General, and a change of numeric codes to mnemonics. More extensive coding was needed to provide for monitoring changes in code values which would not necessarily be communicated to the library between runs of the index program. Departmental name changes, a perennial hazard, result in program notification to the library to check the form of a new name for addition to the library normalization tables, as needed. The chief investigator's name is also inverted for indexing, and checked for alphabetic value.

DURATION OF SYSTEM DESIGN

Elapsed time for project completion was seven weeks, representing five

UCSF	LIBRARY	UCSF CENTRACTS & C	RANTS	INDEX 9/14/73	PAGE
		DEPT. TABLE REPORT			
	(**0.CS=AUUS TO TH	DEPT. TABLE REPORT	FILE	*C=NAME VARIANTY	
0000405	CHANCELLER'S OFFICE		*0	CHANCELLURS OFFICE	
0000405	CHANCELLER'S OFFICE	(0)	*0	CHANCELLOR'S OFFICE	
3000809	-	CHANCELLGR'S CFFICE	*D	CTHER	-
COORDCT	AMB & CCPMUNITY MED	SCHOOL OF PEDICINE	•0	APR & COMM MED	
1008301	BIOCHEM & RIOPHYSICS	SCHOOL OF PEDICINE	*0	BEOCHEN & BECPHYS	
1008001	CARUIGVASCULAR RES	SCHULL OF MEDICINE "	··· *v	CARDIOVAS RES INST	
0008761	PATHCLOGY, CLINICAL	SCHOOL OF MEDICINE	*0	CLIN PATH & LAB MED	
2008901	HISTORY HEALTH SCI	SCHOOL OF MEDICINE	*0	HIST OF HEALTH SCI	
2009101	INTERNATIONAL HEALTH	SCHOOL OF PECICINE	#D-	INTERNAT'L HEALTH	
~~	$\sim\sim$	$\sim\sim\sim$	~	~~~~	~~~~~
110	SE LIABARY	UESF CONTRACTS & G	RANTS_	INDEX 9/14/71	PAGE 2
	ILEZ READ ERRORS	AGENCY CODES NOT, IN	CEG TA	BLE	
		DEPT. CEDES NOT FOUN	D	*	
					100
		PROJ CODES NOT FOUND		SITE CODES NOT FOUND	700 1 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2 1 2
***F[L]	EZ READ STATISTICS				
		NO. OF FILEZ RECORDS	READ:	650	*
		NG. OF FILEZ RECORDS	-Awar	0: 321	
		NC. OF FILEZ RECORDS	-PROPO	SAL: 228	
		NO. OF FILE2 RECORDS	REJEC	TED: 101	
		TOTAL RECORDS PROCES	SED -	650	
		FILEZ TRAILER RECORD	CCUNT	: 000543	

Fig. 5. Error Reports

man-weeks of library staff time and less than a week's time of computer center staff. Total CPU time used was two hours, a significant portion of this time representing compiler rather than program testing. The C&G administrative analyst reviewed initial output and error reports for errors which could have resulted from misunderstanding of C&G format specifications or procedures.

SUMMARY

Open communication channels and interdepartmental cooperation provided the source and means for producing a library information tool which could not have been produced by use of a single department's resources. The data base of local grant information is maintained by a campus administrative department and lent to the library for index extraction. Project development time was considerably shortened by the donated time of computer center staff. The analysis and specification of requirements, and system implementation were provided by the library. Departmental "perquisites" which had to be modified to allow a common project included the library's acceptance of unrevised "bibliographic" information (i.e., unedited names and titles), and the C&G Office's release of an internal file for external use and interpretation.

Work has been completed to provide for inclusion of intramurally funded research projects in the Index. Data for these projects are provided by simplified card input records transcribed by the management services unit of the computer center from routinely generated transmission letters. The resulting index records remain completely independent of the C&G update system except for the use of campus "granting agency" codes added to

the C&G table files.

ACKNOWLEDGMENTS

The virtual absence in this project of the usual irritations and energy-devouring aspects of cooperative work were due to Mardie Petrakis, administrative analyst, and Stanley Bateman, assistant to the vice-chancellor, Academic Affairs, at the UCSF Contracts & Grants Office; Judy Dodge, management systems programmer, Richard Karpinski, systems programmer, and Dr. John Starkweather, director, at the UCSF computer center; Dawne Bissell, special projects assistant, and Jeanette G. Yeazell, university librarian, at the UCSF library. The suggestions and encouragement of the library's reference staff are also gratefully acknowledged.