

File
Copy

INFORMATION RETRIEVAL: APPLICATIONS TO BIBLIOGRAPHIES
ON INTERNATIONAL AND COMPARATIVE POLITICS

Kenneth Janda
Northwestern University

Prepared for delivery at the Computers and the Policy Making Community

Institute: University of California, Lawrence Radiation Laboratory;

Livermore, California; April 4-15, 1966

The phrase "information retrieval" has been applied to many kinds of information processing activities. Although "document retrieval" is a more appropriate label for information processing activities that help locate helpful bibliographic references, both terms will be used more or less interchangeably within this paper. Several information or document retrieval techniques will be discussed with reference to their applications to literature on international relations and comparative politics. The orientation of the paper will be on concrete applications rather than technical considerations.¹ Moreover, its scope will be limited to methods of information retrieval which involve the use of computers, although other techniques and machinery have been used to solve problems of document retrieval.²

The discussion of computerized techniques for document retrieval will be organized in three parts. Part I will discuss automatic indexing of bibliographic references by means of the "keywords" in their titles. Part II will treat in greater detail the Universal Reference System's exhaustive keyword index to selected literature on international affairs. Part III will explain "TRIAL" and "SDI" -- two document retrieval services designed to serve individual research needs.

PART I: KEYWORD INDEXING

Keyword indexing operates on the assumption that certain terms -- "keywords" -- constitute useful handles for locating information of interest. An ordinary book index is an example of a "keyword index." It consists of an alphabetized list of important things discussed in the book along with corresponding page re-

ferences. Most book indexes are prepared by hand, but they readily lend themselves to keyword indexing by computer.³

Preparing an index involves many mechanical tasks: scanning pages for important terms, recording page numbers, cross-listing, and alphabetizing. Familiarity with the subject is undoubtedly an asset in preparing a good index, but a satisfactory job can be expected from any thorough, accurate, and efficient research assistant who is instructed in advance about terms to be included in the index. Computers, which are notoriously thorough, accurate, and efficient, possess the requisite virtues for routinized keyword indexing-- but they require exact instructions in how to do the job.

Computer programs providing these instructions have been available now for some time.⁴ With the proper program, a computer can be instructed to read natural language material, search for the appearance of pre-defined keywords, alphabetize the keywords, and print out the alphabetized keywords along with some of the "context" in which they appear. Two basic versions of these keyword indexing computer programs have become standard information retrieval techniques. One version prints the keywords embedded in the original context; this method is commonly called "KWIC" indexing-- for "Key-Word-In-Context." The other prints the keywords alongside the original context and has become known as "KWOC" indexing-- for "Key-Word-Out-of Context." Both KWIC and KWOC indexing will be discussed in turn.

KWIC Indexing⁵

The first use of KWIC indexing appears to have been in 1958, when IBM published a KWIC index to literature on information retrieval and machine translation.⁶ Within a few years, KWIC indexing was used for preparing bibliographies of recent literature in the physical sciences. A computerized index to Chemical Titles appeared in 1960, and Biological Abstracts mechanized its listing of titles in 1961. One of the earliest, if not the first, applications of KWIC indexing to literature with relevance to political science was in Julian Feldman's bibliography of "Selected Articles and Documents on Methodology and Research in the Social Sciences" published in the September 1960 issue of the American Political Science Review. Two years later, the American Bar Foundation computerized its Index to Legal Theses and Research Projects.⁷ More recently, the

American Political Science Association sponsored the Preparation of a KWIC cumulative index to the first fifty-seven volumes of the American Political Science Review.⁸ The nature of KWIC indexing will be discussed with reference to the cumulative index prepared from articles in the American Political Science Review.

A total of 2,614 articles had appeared in the Review from its first issue in 1906 through its 57th volume in 1963. These articles were punched on IBM cards as illustrated in Figure 1, with different "classes" of cards reserved for different types of information--author, title, and facts of publication.⁹

Figure 1

Several thousand cards carrying citations for all 2,614 articles were fed into an IBM 709 computer. After processing these cards according to program instructions,¹⁰ the computer produced three types of output: an alphabetical listing of all keywords found in the titles of the articles along with part or all of the titles (Figure 2), and alphabetical listing of complete citations by first-named author (Figure 3), and an alphabetical cross-listing of all authors, senior and junior (Figure 4).

Figures 2, 3, and 4

In Figure 2, the keywords are arranged in alphabetical order to the immediate right of the blank column. A keyword is identified by the computer in one of two ways. The computer can be instructed either to refer to a list of previously identified keywords, or to refer to a list of words that are not to be considered keywords. In the first case, the computer looks at every word contained in the title of the publication and compares each one with its own stored list of keywords. The words in the titles found on the keyword list are then selected for indexing. The process operates in a comparable way when a list of non-keywords is used: the computer selects a word for indexing only if the word does not appear in the non-keyword list. A 418 non-keyword list was used in preparing the cumulative index to the Review. Sample non-keywords (words not indexes) were "an," "of," "the," and "other."

The portion of the cumulative index printed in Figure 2 begins with the keyword "FISCAL" and goes to "FRANCE." Between FISCAL and FRANCE, there are

fifty mentions of the keyword "FOREIGN". The fifty titles containing FOREIGN that have been published in the Review since 1906 are arranged in the index according to the alphabetical order of the words which follow FOREIGN.¹¹ The entries, therefore, go from "FOREIGN AFFAIRS" through "FOREIGN SERVICE."

Suppose one were interested in finding references that dealt with "overseas investments" and "foreign trade." While our researcher would want to look under other terms as well, he might begin his search by entering the index under FOREIGN. Scanning down the column, he would find four entries on FOREIGN AID, one on FOREIGN INVESTMENTS, one on FOREIGN LOAN, and another on FOREIGN RAILWAY LOANS. More information about these articles would be obtained by reading the context of the titles surrounding each keyword. There is room on each line for only 78 character-spaces. A title with fewer than 78 characters and spaces prints out in full, although a portion of it may be "wrapped around" and printed before or after the keyword, depending on where the keyword appears in the title.

Suppose that our researcher decided after reading the titles that he was interested only in the two articles dealing with FOREIGN INVESTMENTS AND FOREIGN LOAN. He would determine when these articles were published and learn the author's name by looking at their reference codes in the right-hand column. The reference code gives the first six letters of the senior author's last name, his initials, the year of publication, and the identification number of the article. In our example, we find that both articles were published by W.H. Laves in 1931 and 1934 respectively. The complete citations to these articles can be found under "LAVES WH" in the author-alphabetized bibliography, the relevant portion of which is shown in Figure 3. All of the articles for which Laves is the principal author are given under his name in this author-alphabetized bibliography. The alphabetical cross-listing of authors (Figure 4) shows whether or not Laves appears as a secondary author for any article. We see in Figure 4 that Laves also authored an article with J.W. Fesler in 1951. The complete citation to this article is in the author-alphabetized bibliography under "FESLER JW".

A KWIC index is also known as a "permuted" keyword index, for an article will appear as many times as the number of keywords it contains. The first title

N PENNSYLVANIA.	FISCAL AND ADMINISTRATIVE REORGANIZATION I	KING	CL23	682
STATE SUPERVISION OF LOCAL	FISCAL OFFICERS IN VIRGINIA.	PATE	JE31	1087
LECTED ASPECTS OF INTERGOVERNMENTAL	FISCAL RELATIONS.	KILPAT	W 47	1875
COURT DECISIONS.	REPLY TO FISHER'S 'MATHEMATICAL ANALYSIS OF SUPREME	KORT	F 58	2390
E BRITISH JURISDICTION OVER FOREIGN	FISHERMEN MORE THAN THREE MILES FROM SHORE	GREGOR	CM07	013
NSIBLE PARTIES-- A DISSENT FROM THE	FLOOR.	TURNER	J 51	2059
NE - PARTY LEGISLATIVE SYSTEM-- THE	FLORIDA SENATE, 1947-1961.	PARSON	MR62	2574
EMOCRATIC GUBERNATORIAL PRIMARIES.	FLORIDA'S NEW CAMPAIGN EXPENSE LAW AND THE	ROADY	EE54	2218
11.	GOVERNMENT CORPORATIONS-- A FOCUS OF POLICY AND ADMINISTRATION, I AND	DIMOCK	ME49	1980
AND LEGISLATIVE EFFECTIVENESS.	THE FOLKWAYS OF THE UNITED STATES SENATE-- CON	MATTHE	DR59	2460
D CONSENSUS AMONG PARTY LEADERS AND	FOLLOWERS.	MCCLOS	M 60	2476
AL, STATE, AND LOCAL COOPERATION IN	FOOD AND DRUG CONTROL (LEGISLATION).	CONOVE	M 28	927
TRATION).	FOOD FRONT IN AUSTRALIA (WAR-TIME ADMINIS	DANIEL	M 44	1730
	FOOTSTEPS OF COMMUNITY POWER.	HERSON	LJ61	2539
	IN THE	CONNER	RH49	1956
	UNIFICATION OF THE ARMED	CHEEVE	DS49	1987
N).	FORCES-- THE FIRST YEAR.	MACMAM	AM51	2087
	FOREIGN AFFAIRS I AND THE HOOVER COMMISSIO	STONE	OC52	2147
	FOREIGN AFFAIRS.	MURFEL	D 59	2444
MIC INTEGRATION OF WESTERN EUROPE (FOREIGN AID).	HAVILA	HF58	2429
LOPMENT-- A PHILIPPINE CASE STUDY.	FOREIGN AID AND SOCIAL REFORM IN POLITICAL	MORGEN	M 62	2559
"	FOREIGN AID AND THE POLICY PROCESS-- 1957.	LAVES	WH53	2183
	FOREIGN AID.	JOHNST	WC37	1416
A POLITICAL THEORY OF	FOREIGN ASSIGNMENTS.	DODD	WF11	102
AMERICANS POTENTIALLY AVAILABLE FOR	INCREASING THE NUMBE	DODD	WF10	073
EATY PORTS OF CHINA.	FOREIGN CONCESSIONS AND SETTLEMENTS IN THE	WHITE	LO21	593
CONSTITUTIONAL DEVELOPMENTS IN	FOREIGN COUNTRIES DURING 1910.	RAY	PO18	482
CONSTITUTIONAL DEVELOPMENTS IN	FOREIGN COUNTRIES DURING 1908 AND 1909.	HAINES	CG30	1011
GAMIZATION OF PUBLIC EMPLOYEES (IN	FOREIGN COUNTRIES).	GREGOR	CM07	013
WOMAN SUFFRAGE IN	FOREIGN COUNTRIES.	LAVES	WH31	1073
F JUDICIAL REVIEW OF LEGISLATION IN	FOREIGN COUNTRIES.	LAVES	WH34	1256
AS TO THE BRITISH JURISDICTION OVER	FOREIGN FISHERMEN MORE THAN THREE MILES FR	NIGHTI	RT30	997
TIONAL AND INTERNATIONAL CONTROL OF	FOREIGN INVESTMENTS.	PERKIN	JA40	1522
TOWARD A CONSISTENT	FOREIGN LOAN POLICY.	MORGEN	HJ52	2140
1929.	THE PERSONNEL OF THE BRITISH	HOSKIN	HL53	2161
RS OF INTERNATIONAL IMPORT (MAKING	FOREIGN OFFICE AND DIPLOMATIC SERVICE, 185	NORTHE	F560	2483
NAL INTEREST OF THE UNITED STATES (FOREIGN POLICY).	DENNIS	AL22	643
CURITY PROBLEM IN THE MIDDLE EAST (FOREIGN POLICY).	WILSMA	R 58	2406
	FOREIGN POLICY AND THE PARTY SYSTEM.	CHASE	EP31	1076
EALTH).	BRITISH	TRISKA	JF58	2382
IONAL - EXECUTIVE RELATIONS AND THE	FOREIGN POLICY AND THE DOMINIONS (COMMONW	TAYLOR	P856	2317
	FOREIGN POLICY CONSENSUS.	EPSTEI	LD51	2089
	PARLIAMENTARY CONTROL OF	COLEGR	K 44	1741
	MODEL FOR STUDY OF SOVIET	HOUGHT	ND58	2405
FFAIR-- A CRITIQUE OF UNITED STATES	FOREIGN POLICY.	COOK	T152	2112
TISH LABOUR (PARTY) LEFT AND U.S.	FOREIGN POLICY.	MORGEN	HJ50	2039
S AND PUBLIC OPINION IN FORMULATING	FOREIGN POLICY-- THE REALISM OF IDEALISM.	MARSHA	J 49	1959
TICAL SCIENTISTS IN RECENT AMERICAN	FOREIGN POLICY-- THE NATIONAL INTEREST VS.	EDINGE	LJ63	2608
	FOREIGN POLICY-MAKING).	SCHILL	WR62	2558
TIONS.	THE MAINSPRINGS OF AMERICAN	PRINCE	C 44	1728
CITIZEN DIPLOMACY (DEMOCRATIZING	FOREIGN POLICY-MAKING).	SPYKMA	NJ39	1483
MILITARY LEADERS AND	FOREIGN POLICY-MAKING.	SPYKMA	NJ38	1427
SCIENTISTS,	FOREIGN POLICY, AND POLITICS.	CARROL	EM52	2125
C FACTORS AFFECTING SOVIET RUSSIA'S	FOREIGN POLICY, I AND II.	WANG	CC10	075
GEOGRAPHIC OBJECTIVES IN	FOREIGN POLICY, I AND II.	MYERS	DP17	411
GEOGRAPHY AND	FOREIGN POLICY, I AND II.	MYERS	DP17	446
CENT GERMAN PUBLICATIONS AND GERMAN	FOREIGN POLICY, 1933-1945.	WRIGHT	Q 21	578
WHY THE CHINESE OPPOSE	FOREIGN RAILWAY LOANS.	GRAHAM	HW25	744
"	THE CONTROL OF FOREIGN RELATION (TACTICS OF DIPLOMACY).	SHARP	WR44	1739
"	LEGISLATURES AND FOREIGN RELATIONS (A COMPARATIVE STUDY).	FOREIG	RE58	2401
).	THE CONTROL OF FOREIGN RELATIONS (PRESIDENT AND CONGRESS	FOREIG	RE61	2533
INSTRUCTION, CONSTITUTIONS, PARTIES,	FOREIGN RELATIONS).	LAY	T 24	723
D FRAMEWORK.	AMERICAN	MACCLI	S 22	644
REPORT OF THE ADVISORY COMMITTEE ON	FOREIGN RELATIONS TO THE DEPARTMENT OF STA	MCCANY	JL54	2239
REPORT OF THE ADVISORY COMMITTEE ON	FOREIGN RELATIONS-- 1960 .	STEWAR	I 30	999
	FOREIGN SERVICE REORGANIZATION.	LAWSON	MG57	2379
	FOREIGN SERVICE.	HASTIN	AB30	1003
PEOPLE OF THE STATE DEPARTMENT AND	FOREIGN SERVICE.	PEYTON	J529	943
TE.	CONGRESS, THE	LASSWE	MD25	762
ISTICAL SUMMARY.	THE FOREIGN-BORN IN CONGRESS, 1789-1949-- A ST	BATES	FG10	167
	PROGRESS IN STATE	SINGER	JD63	2610
	STATE	SHERMA	RV32	1131
	TWO FORGOTTEN STUDIES IN POLITICAL PSYCHOLOGY.	FIELD	GL33	1201
MUNICIPAL GOVERNMENT-- COMMISSION	FORM.	LOWRIE	SG11	084
INTER-NATION INFLUENCE-- A	FORMAL MODEL (INTERNATIONAL RELATIONS).	WILCOX	FO45	1795
ERTAKINGS (FASCIST ECONOMICS).	FORMS OF MUNICIPAL GOVERNMENT IN OHIO.	COLEGR	K 44	1741
	FORMS OF ORGANIZATION OF ITALIAN PUBLIC UN	LANGRO	G 55	2271
	NEW FORMS OF THE INITIATIVE AND REFERENDUM.	MARX	FM39	1474
E OF CONGRESS AND PUBLIC OPINION IN	FORMULA (IN THE U.N.).	SOUSERS	SW49	1973
COUNCIL OF STATE-- ITS ROLE IN THE	FORMULATING FOREIGN POLICY.	BENNET	EL39	1471
).	FORMULATION AND IMPLEMENTATION OF ADMINIST	BEARD	CA40	1519
E).	POLICY FORMULATION AND THE ADMINISTRATIVE PROCESS	ASPATU	VY54	2237
).	POLICY FORMULATION FOR NATIONAL SECURITY (DEFENS	WICKWA	WM47	1899
).	FORUM OF REMOVALS AND RETIREMENTS (IMPEAC	ROCHE	JP61	2538
).	FOUNDATION OF GOVERNMENT.	ANDERS	M 55	2256
SOVIET STATE AND ITS PHILOSOPHICAL	FOUNDATIONS (COMMUNIST THEORY).	DIAMON	M 59	2432
	FOUNDATIONS OF AMERICAN CONSERVATISM.	BACHRA	P 63	2621
E CONSTITUTIONAL CONVENTION).	FOUNDING FATHERS-- A REFORM CAUCUS IN ACTI	SHARP	WR44	1739
ETATION.	FRAMERS-- A NOTE ON CONSTITUTIONAL INTERPR	KOESSL	M 45	1764
DERALIST-- A RECONSIDERATION OF THE	FRAMERS' INTENT.	MACRID	R 56	2321
NS AND NONDECISIONS-- AN ANALYTICAL	FRAMEWORK.			
RELATIONS WITHIN AN ORGANIZED WORLD	FRAMEWORK.			
VICHY'S SHAM CONSTITUTIONALITY I	FRANCE).			
NSTITUTION OF THE FOURTH REPUBLIC I	FRANCE).			

FIGURE 2: KWIC computer printout: keyword index from the Cumulative Index to the American Political Science Review

0296	LATANE JH14	LATANE JH	OUR RELATIONS WITH JAPAN. AM POL SCI REV 8 (NOVEMBER 1914) 503-601
2114	LATHAM E 52	LATHAM E	THE GROUP BASIS OF POLITICS-- NOTES FOR A THEORY. AM POL SCI REV 46 (JUNE 1952) 376-397
1073	LAVES WH31	LAVES WHC	NATIONAL AND INTERNATIONAL CONTROL OF FOREIGN INVESTMENTS. AM POL SCI REV 25 (AUGUST 1931) 704-713
1256	LAVES WH34	LAVES WHC	TOWARD A CONSISTENT FOREIGN LOAN POLICY. AM POL SCI REV 28 (DECEMBER 1934) 1047-1052
1550	LAVES WH40	LAVES WHC	THE NEXT DECADE IN POLITICAL SCIENCE TEACHING. AM POL SCI REV 34 (OCTOBER 1940) 983-986
1696	LAVES WH49	LAVES WHC	THE FACE-TO-FACE WAR INFORMATION SERVICE OF THE FEDERAL GOVERNMENT. AM POL SCI REV 37 (DECEMBER 1943) 1027-1040
1712	LAVES WH44	LAVES WHC	THE REORGANIZATION OF THE DEPARTMENT OF STATE. AM POL SCI REV 38 (APRIL 1944) 289-301
1738	LAVES WH44	LAVES WHC	ORGANIZING THE GOVERNMENT FOR PARTICIPATION IN WORLD AFFAIRS (INTERNATIONAL RELATIONS). AM POL SCI REV 38 (OCTOBER 1944) 913-930
1770	LAVES WH49	LAVES WHC	THE STATE DEPARTMENT CONTINUES ITS REORGANIZATION. AM POL SCI REV 39 (APRIL 1945) 309-317
2103	LAVES WH53	LAVES WHC	INCREASING THE NUMBER OF AMERICANS POTENTIALLY AVAILABLE FOR FOREIGN ASSIGNMENTS. AM POL SCI REV 47 (SEPTEMBER 1953) 798-804
925	LAWREN D 28	LAWRENCE D	REPORTING THE POLITICAL NEWS AT WASHINGTON. AM POL SCI REV 22 (NOVEMBER 1928) 893-902
2379	LAWSON MG57	LAWSON MG	THE FOREIGN-BORN IN CONGRESS, 1789-1949-- A STATISTICAL SUMMARY. AM POL SCI REV 51 (DECEMBER 1957) 1183-1189
723	LAY T 24	LAY T	FOREIGN SERVICE REORGANIZATION. AM POL SCI REV 18 (NOVEMBER 1924) 697-711

FIGURE 3: KWIC computer printout: author-alphabetized bibliography from the Cumulative Index to the American Political Science Review

LAVES WHC	LAVES WH31	LOWENSTEIN K	LOWENS K 37
LAVES WH34	LAVES WH40	LOWENTHAL D	LOWENS K 44
LAVES WH43	LAVES WH44	LOWI T	LOWENT D 59
LAVES WH44	LAVES WH45	LOWRIE SG	LOWENT D 63
LAVES WH45	FESLER JWS1		LOWI T 63
LAVES WH53	LAVES WH53		LOWRIE SG11
LAWRENCE D	LAWREN D 28		LOWRIE SG11
LAWSON MG	LAWSON MG57		LOWRIE SG11
LAY M	MYNEMA CS38		LOWRIE SG11
LAY T	LAY T 24		LOWRIE SG11
LAYLIN CD	LAYLIN CD23		LOWRIE SG13
LEACOLK S	LEACOC S 07		LOWRIE SG22
	LEACOC S 10		LOWRIE SG26
LEAKE JM	LEAKE JM17		LOWRIE SG27
	LEAKE JM18		LOWRIE SG28
LEARNED MB	LEARNE HB09		LOWRIE SG36
	LEARNE HB21		LOWRIE SG51
LEDERLE JM	LEDERL JH46	LUCE R	LUCE R 32
LEE FP	LEE FP28	LUDINGTON A	LUDING A 11
LEEK JM	LEEK JM26		LUDING A 09

FIGURE 4: KWIC computer printout: author cross-listing from the Cumulative Index to the American Political Science Review

in our example, "National and International Control of Foreign Investments," will be listed in the index under "NATIONAL," "INTERNATIONAL," "CONTROL," and "INVESTMENTS". All these terms are identified as keywords by the computer because they do not appear on the non-keyword list. The permuting feature of the KWIC index produced 10,089 keyword lines from the titles of 2,614 articles in the Review. Therefore each title appears in the index on an average of 3.9 times.

KWOC Indexing

Conventional KWIC indexes are "double-entry" indexes; after the user had found a title of interest, he must use the reference code to enter the author-alphabetized bibliography for complete information on journal, volume, month, year, and pages. Although it would be possible to develop a KWIC index that would give this information in abbreviated form in the reference code, the purpose of a "single-entry" index is perhaps better served by the KWOC version of keyword indexing, which prints out the complete citation for each appearance of a keyword in the title.

KWOC indexing came as a refinement of KWIC indexing and had not been as widely used.¹² An illustration of KWOC indexing will be provided by an unpublished index to 928 titles on "Africa" and "The Middle East" as reported in the "Foreign and Comparative Government" bibliographies of the American Political Science Review from 1959 to 1963.¹³ Figure 5 reproduces a page of computer output from this index for keywords beginning with "COMMUNISM". The appearance of a keyword in each title is indicated by underlinings.

Figure 5

The punchcard format for input to the KWOC program is identical to the KWIC format (Figure 1), but the computer is instructed to treat the input differently.¹⁴ Both KWIC and KWOC programming instructions identify the keywords or non-keywords. But when a keyword is found under the KWOC program, the computer reprints it out of context to the left of the complete citation, which includes the author, title, and sources. KWIC, by printing only one line per entry, is definitely more economical in its use of space and is possibly easier to use in locating titles of interest. KWOC, on the other hand, provides a single-entry look-up and allows for printing complete titles regardless of length.

It appears that KWIC indexing is more suitable for preparing cumulative indexes to articles from the same journal, like the cumulative index to the Review. Because all articles are from the Review, the user need not turn to the author-alphabetized listing to determine where the article had been published. If desired, the reference code could be expanded to include the month of publication as well as the year, which would further reduce the need to consult the author-alphabetized bibliography. KWOC, on the other hand seems to be more suitable for indexing titles taken from diverse sources, as illustrated in Figure 5, and for porcessing titles extending over two or more lines.¹⁵

Advantages of Keyword Indexing

The advantages of automatic indexing of bibliographical material by keywords in titles are those generally associated with the use of computers in data processing. First of all, the indexes are easily and inexpensively prepared. The input to the KWIC index of the Review was prepared by a girl keypunching directly from title pages of bound library volumes. It took her less than 200 hours to punch and correct all the titles in 57 volumes of the Review. It took the IBM 709 computer less than 12 minutes to process all 2,614 titles -- which meant searching a 418 non-keyword list for each word in every title, preparing 10,089 KWIC index lines, and producing a cross-reference listing for 2,801 senior and junior authors. Another 29 minutes were required to sort the output into alphabetical order, and about 30 minutes were needed to print the output on the IBM 1401. Similar economics were realized in preparing the KWOC index to titles on Africa and the Middle East. This 928 item bibliography was prepared in less than two weeks for less than \$150.

In addition to the advantages of speed and economy, computer-generated keyword indexes are easily updated with new material and are readily reproduced. Once a comprehensive bibliography is punched on cards, it becomes a simple matter to prepare specialized bibliographies by instructing the computer to index the literature on a smaller number of previously identified keywords, such as "conflict," "aggression," "war," "revolution," "attack," etc. And once the bibliography is punched in machine-readable form, the computer can retrieve references according to more complex search commands-- the subject of Part III.

469	COMMUNISM	UDOCI JM	THE CONFLICT INVOLVING <u>COMMUNISM</u> IN MID- AFRICA.= THE ANNALS JULY, 1962
790	COMMUNISM	MORISON DL	<u>COMMUNISM</u> IN AFRICA -- MOSCOW'S FIRST STEPS.= PROBLEMS OF COMMUNISM NOVEMBER-DECEMBER, 1961
791	COMMUNISM	KOLARZ W	<u>COMMUNISM</u> IN AFRICA -- THE WEST AFRICAN SCENE.= PROBLEMS OF COMMUNISM NOVEMBER-DECEMBER, 1961
2027	COMMUNISM	LAQUEUR WZ	<u>COMMUNISM</u> AND NATIONALISM IN TROPICAL AFRICA.= FOREIGN AFFAIRS JULY, 1961
1320	COMMUNIST	AUTHOR NOT GIVEN	THE <u>COMMUNIST</u> PARTY IN SOUTH AFRICA.= AFRICA REPORT MARCH, 1961
540	COMMUNITY	SAFRAN N	EGYPT IN SEARCH OF POLITICAL <u>COMMUNITY</u> -- AN ANALYSIS OF THE INTELLECTUAL AND POLITICAL EVOLUTION OF EGYPT 1804-1952.= CAMBRIDGE -- HARVARD UNIVERSITY PRESS, 1961 (REVIEWED IN AM POL SCI REV 56, SEPTEMBER, 1962 PP748)
971	COMMUNITY	IRVINE K	THE FRANCO- AFRICAN <u>COMMUNITY</u> IN TRANSITION.= CURRENT HISTORY OCTOBER, 1961
2370	COMMUNITY	RUDLIN HR	GUINEA OUTSIDE THE FRENCH <u>COMMUNITY</u> .= CURRENT HISTORY (JULY 1959)
075	COMPARATIVE	ABU-LUGHOD I	INTERNATIONAL NEWS IN THE ARABIC PRESS-- A <u>COMPARATIVE</u> CONTENT ANALYSIS.= PUBLIC OPINION QUARTERLY DECEMBER, 1962
A2282	COMPARATIVE	KAHIN GMCT PAUKER GJ PYE LW	<u>COMPARATIVE</u> POLITICS OF NON-WESTERN COUNTRIES.= AM POL SCI REV 49 (DECEMBER 1955) 1022-1041
A2354	COMPARATIVE	BINDER L	PROLEGOMENA TO THE <u>COMPARATIVE</u> STUDY OF MIDDLE EAST GOVERNMENTS.= AM POL SCI REV 51 (SEPTEMBER 1957) 651-668
2349	COMPETE	GORDON W	IN NEWEST AFRICA-- THREE NATIONALISMS <u>COMPETE</u> FOR THE EMERGING PEOPLES.= NIEMAN REPORTS (JULY 1959)
1746	COMPOSITION	JEAREY JH	THE STRUCTURE, <u>COMPOSITION</u> AND JURISDICTION OF COURTS AND AUTHORITIES ENFORCING THE CRIMINAL LAW IN BRITISH AFRICAN TERRITORIES.= INTERNATIONAL AND COMPARATIVE LAW QUARTERLY JULY, 1960
A2598	CONCEPT	MAZRUI AA	ON THE <u>CONCEPT</u> OF 'WE ARE ALL AFRICANS'.= AM POL SCI REV 57 (MARCH 1963) 88-97

FIGURE 5: KWOC computer printout: keyword index from 928 titles on
politics in Africa and the Middle East

Disadvantages of Keyword Indexing

Obviously the major disadvantage in indexing literature by keywords in titles lies in the "descriptiveness" of titles fed into the computer. Keyword indexing was originally developed for "technical" literature, and indeed the longer and more descriptive titles of articles in the physical and biological sciences seem better suited to this technique than titles in the behavioral sciences. Lane has investigated variation within different fields for the suitability of indexing titles solely by their keywords.¹⁶ He points out that the reader interested in "duck shooting" probably would have missed an article entitled "Good Day in Bad Marsh" if he had consulted a keyword index, but the article would have been found under "duck shooting" in the subject heading index of the Readers' Guide to Periodical Literature.

Evidence indicates, however, that scholarly literature in social science has a higher proportion of "descriptive" titles than the largely popular literature reported in the Readers' Guide.¹⁷ Moreover, the problem of inadequate titles can be partially solved at least through editorial supervision while the computer input is being prepared. Scanning the texts of articles whose titles seem unclear or literary in nature will usually disclose some terms or phrases which might be enclosed in parentheses and added to the titles. Keywords added in this manner will be indexed as if they had been in the original title. This procedure was followed in preparing titles from the Review for computer processing. For example, the first three FOREIGN POLICY entries in Figure 2 are keyword additions to titles that did not mention "foreign policy," although that is the subject of the articles. In all, keywords were added to 604 of the 2,614 Review titles, indicating that original titles were considered suitable "indexable" in about 77 percent of the cases.

Effective editorial supervision is relatively easy to exercise when the complete article is available for examination, as it is while punching directly from journals. But when preparing computer input from bibliographical citations, there is little or no opportunity to examine the sources cited. In this case, indexing effectiveness depends completely on the information contained in the titles: the less descriptive the titles, the more imperfect the index. When judged by some absolute standard of "good" indexing, an unedited keyword index to political

science titles is likely to look very bad indeed. When judged according to available indexes, however, even an unedited keyword index often looks good. Keyword indexing is so quick, easy, and inexpensive that computerized indexes can be produced when handcrafted ones are not justifiable. Granting imperfections in the product, a keyword index is virtually always preferable to no index at all. It may miss many references, but it will also retrieve many that otherwise would escape notice.

Clearly, however, the best way to improve the completeness of keyword indexes is to provide more information about the publications in the input to the computer. Short of keypunching the entire article, each citation could be tagged with an extensive set of descriptive terms abstracted from the item itself. This approach was used to create the comprehensive index to International Affairs, Volume I of the Universal Reference System.

PART II: UNIVERSAL REFERENCE SYSTEM

The Universal Reference System (URS) was developed under the direction of Professor Alfred de Grazia, editor of the American Behavioral Scientist. The URS had its beginnings in the annotated bibliographies that have been a regular feature of the American Behavioral Scientist and its predecessor, Political Research: Organization and Design.¹⁸ In 1963, de Grazia undertook plans to computerize the bibliographical items assembled over the years. His plans developed into the URS "Political Science, Government & Public Policy Series," which is scheduled to consist of ten volumes on (1) international affairs, (2) legislative processes, (3) administrative management, (4) current events, (5) public opinion, (6) law, (7) economic regulation, (8) public policy, (9) comparative government, and (10) bibliographies in political science. The first volume of this Series, International Affairs, appeared in 1965.¹⁹

International Affairs indexes some 3,030 books, articles, and pamphlets on international affairs selected from material "above a modest level of quality, appearing in all social and behavioral sciences."²⁰ Entries are drawn primarily from literature published from 1962 to 1965, but coverage extends to World War I and earlier to include such classic studies as Machiavelli's The Prince. Each entry is annotated and tagged with "Standard" and "Unique" descriptors. The descriptors seek to describe the entry in terms of "Who says, 'Who does what with

whom, where and when, by what means, why' and how does he know so?"²¹

The URS classificational system contains some 183 Standard Descriptors organized in a topical and methodological index according to the following outline.²²

I. TOPICS (Action pattern described in the work)

1. TIME-SPACE-CULTURE INDEX: Where ?
2. INSTITUTIONAL INDEX (TYPICAL ROLE): Who? Whom?
3. ORGANIC INDEX: By what?
4. SCOPE INDEX (Dimensions of Aggregates Treated)
5. PERSONALITY INDEX: In what characterization?
6. VALUES INDEX: For what goals?
7. MANIPULATIVE TACTICS INDEX: By what tactics?

II. METHODOLOGY (Used or dealt with by author)

8. ETHICAL STANDARDS APPLIED
9. ETHICAL BASE INDEX
10. FIELD INDEX
11. CONCEPTS
12. LOGIC, MATHEMATICS AND LANGUAGE
13. DIRECT OBSERVATION
14. INTERVIEWS
15. QUESTIONNAIRES
16. TESTS AND SCALES
17. UNIVERSES AND SAMPLING
18. ANALYSIS OF TEMPORAL SEQUENCES
19. COMMUNICATION CONTENT ANALYSIS
20. INFORMATION STORAGE AND RETRIEVAL
21. GRAPHICS AND AUDIO-VISUAL TECHNIQUES
22. EXPERIMENTATION
23. MODELS
24. GENERAL THEORY

In order to save space in the computer printout, most of these descriptors have been truncated into code words -- prompting the volume to be called "CODEX" instead of "index." To understand the notion of "truncated descriptors," consider those under heading #20, INFORMATION STORAGE AND RETRIEVAL:

AC181-A0193

EXAMINATION OF THREE ECONOMIC SYSTEMS AS ALTERNATIVE AND COMPETITIVE MODELS FOR THE ECONOMIC DEVELOPMENT OF UNDERDEVELOPED COUNTRIES. PARTICULARLY OF THOSE IN ASIA. SUGGESTS THAT THE INDIAN EXAMPLE OF HIGH PRIORITY TO WELFARE GOALS AND LIMITATIONS ON MARKET FORCES BY PARTIAL PLANNING AND SOME DIRECT CONTROLS MIGHT BE FOLLOWED.

CHINJAP S/ASIA EC/UNDEV ECO/DEV MARKET WEALTH...ECO DOC/ANAL.

AC182 NORTH R.C. ET AL.

CONTENT ANALYSIS: A HANDBOOK WITH APPLICATIONS FOR THE STUDY OF INTERNATIONAL CRISIS.

EVANSTON: NORTHWESTERN U. PR., 1963. 182 PP., \$2.95.

A VALUABLE INTRODUCTION TO A USEFUL RESEARCH TECHNIQUE. INCLUDES CONCRETE ILLUSTRATIONS PLUS GUIDES 'FOR DECIDING WHETHER, WHEN, AND WHAT FORM OF CONTENT ANALYSIS SHOULD BE USED.' THE EXAMPLES COME FROM RESEARCH ON THE ORIGINS OF WORLD WAR I AND ON CONTEMPORARY SINO-SOVIET RELATIONS. SPECIAL FORMS DISCUSSED INCLUDE: THE CONVENTIONAL FREQUENCY COUNT AND QUALITATIVE IDENTIFICATIONS, Q-SCATS, PAIR COMPARISONS, AND EVALUATIVE ASSERTION ANALYSIS.

CHINJAP COM EUR+WW1 MOD/EUR USSR INT/ORG ROUTINE TEC/DEV COMMUN PERCEPT RIGID/FLEX ALL/VALS COERCE DOMIN EDU/PROP... POL QUANT TESTS COM/ANAL COMPUT/IR SIMUL GEN/LAWS GEN/METH.

AC183 LERCHE C.O. JR., SAID A.A.

CONCEPTS OF INTERNATIONAL POLITICS.

ENGLEWOOD CLIFFS: PRENTICE-HALL, 1963. 314 PP., \$4.95.

SYSTEMATIC DEVELOPMENT OF THE CONCEPTS OF THE STATE AND FOREIGN POLICY, THE NOTION OF POLITICS AS THE CRUCIAL FORM OF INTERSTATE RELATIONSHIPS, AND SUBSTANTIVE PROBLEMS OF INTERNATIONAL POLITICS SUCH AS WAR, IDEOLOGY, TECHNOLOGY, TRADE, AND COMMUNICATION.

WOR+45 WOR+45 INT/ORG INT/TRADE DELIB/GP EX/STRUC LAW ROUTINE TEC/DEV ATTIT-ORD/FREE PWR RESPECT COERCE ECO/TAC LEGIT WAR...POL DOC/ANAL STENTYP GEN/LAWS.

A0184 ARDANT G.

A PLAN FOR FULL EMPLOYMENT IN THE DEVELOPING COUNTRIES.

INT. LABOUR REV., 87 (JAN. 63), 15-51.

OUTLINES A PLAN FOR PRODUCTIVE EMPLOYMENT OF UNEMPLOYED AND UNDEREMPLOYED MASSES OF DEVELOPING COUNTRIES, BASED LARGELY ON PRINCIPLE OF SELF-HELP AND MAKING RELATIVELY MODEST DEMANDS ON FINANCIAL OR SKILLED-MANPOWER RESOURCES. CONCLUSIONS ARE DERIVED MAINLY FROM THE AUTHOR'S EXPERIENCE WITH EMPLOYMENT PROJECTS IN MOROCCO AND TUNISIA.

FUT MOROCCO TUNIS WOR+45 EC/UNDEV SOCIETY ACT/RES DELIB/GP PLAN ATTIT ALL/VALS ECO/TAC...POLICY ECO PHIL/SCI STAT CHARTS.

A0185 BRAIBANTI R.J.D., SPENGLER J.J.

ADMINISTRATION AND ECONOMIC DEVELOPMENT IN INDIA.

DURHAM: DUKE U. PR., 1963. 312 PP., \$7.00.

INDIAN AND AMERICAN POLITICAL SCIENTISTS AND ECONOMISTS DISCUSS THE RELATIONSHIP OF ECONOMIC DEVELOPMENT AND ADMINISTRATION IN INDIA. ECONOMIC DEVELOPMENT 'HAS BEEN BLOCKED TO A LARGE EXTENT PRECISELY BECAUSE ITS IMPLEMENTATION DEPENDS ON A STRUCTURE AND DISPOSITION OF BUREAUCRACY, NEITHER OF WHICH CAN BE JARRED OUT OF THE LARGER SOCIETAL WHOLE.'

S/ASIA EC/UNDEV SOCIETY ADMIN STRATA PERSON WEALTH ECO/TAC ...ECO NGT GEN/LAWS.

PUBLIC ADMINISTRATION OFFER A MORE BALANCED VIEW OF ECONOMIC DEVELOPMENT PROBLEMS IN LATIN AMERICA.

L/A+17C EC/UNDEV FINAN INDUS INT/ORG SOCIETY ACT/RES ADMIN CULTURE DELIB/GP STRATA ATTIT SUPEGO HEALTH KNOWL ORD/FREE ECO/TAC EDU/PROP...ECO SOC STAT TREND DOC/ANAL ANTHOL.

A0188 MAIZELS A.

INDUSTRIAL GROWTH AND WORLD TRADE.

NEW YORK: CAMBRIDGE U. PR., 1963. 563 PP., \$14.00.

'AN EMPIRICAL STUDY OF TRENDS IN PRODUCTION, CONSUMPTION AND TRADE IN MANUFACTURES FROM 1899-1959 WITH A DISCUSSION OF PROBABLE FUTURE TRENDS.' ATTEMPTS TO ANSWER THE IMPORTANT QUESTION OF WHETHER INDUSTRIALIZATION OF LESS-DEVELOPED COUNTRIES WOULD ULTIMATELY RESULT IN AN EXPANSION OR A CONTRACTION IN DEMAND FOR MANUFACTURED EXPORTS OF INDUSTRIAL COUNTRIES.

FUT WOR+45 EC/UNDEV ECO/DEV FINAN INDUS INT/ORG INT/TRADE PLAN TEC/DEV WEALTH ECO/TAC...ECO MATH STAT DOC/ANAL CHARTS.

A0189 WILCOX F.O. EC., HAVILAND H.F.

THE ATLANTIC COMMUNITY: PROGRESS AND PROSPECTS.

INT. ORGAN., 17 (SUMMER 63), 521-529.

EXAMINES THE QUESTION OF WHETHER COOPERATIVE RELATIONSHIPS AMONG THE ATLANTIC NATIONS SHOULD BE DEVELOPED FURTHER, AND IF SO, FOR WHAT ENDS AND BY WHAT MEANS. INCLUDES ARTICLES BY STANLEY JOFFMAN, 'DISCORD AND COMMUNITY: THE NORTH ATLANTIC AREA AS A PARTIAL INTERNATIONAL SYSTEM', RUPERT EMERSON, 'THE ATLANTIC COMMUNITY AND WESTERN INTEGRATION', AND H. F. HAVILAND, JR., 'BUILDING A POLITICAL COMMUNITY.'

EUR+WW1 FUT USA+45 WOR+45 INT/ORG REGION SOCIETY ACT/RES CREATE ATTIT ALL/VALS ECO/TAC EDU/PROP LEGIT...POLICY ECO POL ANTHOL.

A0190 KUZNETS S.

QUANTITATIVE ASPECTS OF THE ECONOMIC GROWTH OF NATIONS: DISTRIBUTION OF INCOME BY SIZE.

ECON. DEVELOP. CULT. CHANGE, 11 (JAN. 63), 1-80.

AN EXTENSIVE STUDY DEALING WITH RELATION BETWEEN ECONOMIC GROWTH AND DISTRIBUTION OF INCOME BY SIZE AMONG INDIVIDUALS AND HOUSEHOLDS IN A COUNTRY. ATTEMPTS TO ANSWER TWO CRUCIAL QUESTIONS: FIRST, DOES ECONOMIC DEVELOPMENT AFFECT DISTRIBUTION OF GROWING INCOME AMONG THE POPULATION. SECOND, DOES THE SIZE DISTRIBUTION OF INCOME, AFFECTED BY ECONOMIC GROWTH, HAVE IN TURN AN EFFECT ON THE LATTER.

WOR+45 FINAN ACT/RES HEALTH WEALTH ECO/TAC...ECO MATH STAT.

A0191 ROSE R. EC., FEIDENHEIMER A. ED.

COMPARATIVE STUDIES IN POLITICAL FINANCE: A SYMPOSIUM.

J. POLIT., 25 (NOV. 63), 643-671.

A COLLECTION OF PAPERS BY WELL-KNOWN SPECIALISTS. EXAMINING INTENSIVELY THE PARTICULAR PATTERN OF POLITICAL FINANCE IN EIGHT DEMOCRATIC STATES. EACH ARTICLE CONCERNS A BASIC SET OF QUESTIONS ABOUT PARTY MEMBERSHIP, INCOME, EXPENDITURE, SUBSIDIES, LAWS AND CAMPAIGN COSTS. INCLUDES AUSTRALIA, BRITAIN, INDIA, ISRAEL, ITALY, JAPAN, THE PHILIPPINES, AND WEST GERMANY.

AUSTRAL BRIT CHINJAP EUR+WW1 GERMANY ISRAEL PHIL S/ASIA FINAN POL/PAR ACT/RES CHOICE CULTURE DELIB/GP LAW LEGIS ATTIT COMMUN RIGID/FLEX SUPEGO PWR SKILL WEALTH ECO/TAC EDU/PROP...ECO POL STAT DOC/ANAL ANTHOL.

FIGURE 6: CODEX computer printout: Catalog of items from International Affairs, Volume I of the Universal Reference System, page 14.

COMMUNIST COUNTRIES (EXCEPT CHINA)....SEE COM

COMPANY, LARGE OR MIDDLE-SIZED (50 EMPLOYEES OR MORE)....
SEE LG/CO

COMPANY, SMALL (50 EMPLOYEES OR LESS)....SEE SML/CO

COMPETENCE, PRACTICAL (AS GOAL)....SEE SKILL

COMPUT/IR....MECHANICAL AND ELECTRONIC DEVELOPMENTS FOR INFOR-
MATION RETRIEVAL.

COMPUT/IR B=51	ATTIT SAMP	WOR+45 WOR-45 COM/IND R+D SOCIETY ACT/RES CHOICE CREATE DELIB/GP PERSON KNOWL....STAT DEEP/INT STAND/INT SAMP/SIZ CHARTS STERTYP	A1322
COMPUT/IR B=52	PERSON PSY	FUT WOR+45 CULTURE COMMUN PERCEPT.... CONT/OBS REC/INT PERS/TEST SAMP LAB/EXP GEN/METH	A1023
COMPUT/IR B=58	TEC/DEV SOC CAP/ISM	FUT USA+45 USA-45 ECO/DEV LABOR SOCIETY ACT/RES PLAN ROUTINE WORK ALL/VALS ECO/TAC....ECO STAT TREND CHARTS SOC/EXP	A1532
COMPUT/IR B=59	COM STRUCT USSR	MOD/EUR ECO/DEV SOCIETY PLAN ATTIT SOC/ISM COERCE GORD/WAR ECO/TAC EDU/PROP....ART/METH ECO HIST SOC CONCEPT STAT DOC/ANAL GEN/LAWS GEN/METH ANTHOL	A0805
COMPUT/IR B=60	INT/ORG JURID SPACE INT/LAW	FUT WOR+45 ADJUD CONSTN CONSULT PLAN STRUCT TEC/DEV RIGID/FLEX KNOWL ORD/FREE PWR....TECHNIC GEOG MGT NEW/IDEA TREND DOC/ANAL	A0457
COMPUT/IR B=60	MARKET ADJUD EEC INT/LAW	EUR+WW1 FUT ECO/DEV INT/TRADE NAT/G SOCIETY VOL/ASSN ADMIN CONSULT PLAN STRUCT TEC/DEV ATTIT RIGID/FLEX PWR ECO/TAC....ECO MGT NEW/IDEA STAT TREND SIMUL	A0591
COMPUT/IR B=63	WOR+45 POL/PAR	COM/IND EC/UNDEV ECO/DEV NAT/G SECT CHOICE CONSTN CT/SYS EX/STRUC FORCES LEGIS TOP/EX STRATA PWR HEALTH....POL STAT CON/ANAL CHARTS SIMUL	A0147
COMPUT/IR B=63	COMPUT/IR USSR	CHINJAP COM EUR+WW1 MOD/EUR INT/ORG ROUTINE TEC/DEV COMMUN PERCEPT RIGID/FLEX ALL/VALS COERCE DOMIN EDU/PROP....POL QUANT TESTS CON/ANAL SIMUL GEN/LAWS GEN/METH	A0182
COMPUT/IR S=59	COMPUT/IR DECISM DIPLOM	FUT WOR+45 NAT/G R+D ACT/RES CHOICE CREATE ATTIT PWR....POLICY POL CONCEPT STYLE KNO/TEST TREND HYPO/EXP GEN/METH	A0610
COMPUT/IR S=63	POL DECISM	FUT UNIV CONSULT PLAN SKILL COERCE DETER DOMIN EDU/PROP WAR....QUANT HYPO/EXP	A1219
COMPUT/IR S=63	PLAN TEC/DEV USSR	COM FUT USA+45 DIST/IND ECO/DEV MARKET RIGID/FLEX CAP/ISM SOC/ISM WEALTH ECO/TAC....MATH STAT TREND DOC/ANAL	A1767
COMPUT/IR S=64	FORCES SIMUL WAR	USA+45 NAT/G ACT/RES TEC/DEV TOP/EX KNOWL ORD/FREE PWR COERCE....PHIL/SCI POL MATH TIME/SEQ CHARTS LAB/EXP DECISM	A0285

CON/ANAL....QUANTITATIVE CONTENT ANALYSIS= OBJECTIVE,
SYSTEMATIC AND QUANTITATIVE DESCRIPTION OF THE MANIFEST
CONTENT OF COMMUNICATION.

FIGURE 7: CODEX computer printout: Index of Documents from International Affairs, Volume I of the Universal Reference System, page 395

THING/STOR	Artifacts and material evidence.
OLD/STOR	Conventional systems: libraries, books, records, tape, film as well as the human brain.
COMPUT/IR	Mechanical and electronic developments: for information retrieval

In addition to the 183 Standard Descriptors, there are 121 Unique Descriptors (e.g., "POLAND") which have been selected for their relevance to international affairs. The Unique Descriptors may or may not be truncated, depending on their length.

The value of the URS international affairs CODEX derives from the in-depth indexing of its entries. Whereas most keyword indexes to bibliographies rely solely on titles (perhaps supplemented with additional keywords), the URS input is not only annotated but is also tagged with an average of twenty Standard and Unique Descriptors per item. Moreover, from two to four of these descriptors are identified as "Critical" Descriptors and are given special treatment in indexing. To illustrate the nature of the punchcard input to URS, some of the 3,030 entries in the CODEX "Catalog" are reproduced in Figure 6.

Figure 6

The program for processing the URS entries and producing the keyword index was written for the IBM 1401/1410 computers.²³ Essentially, the program is a type of keyword indexing, using the descriptors as keywords. The format of the computer printout is similar to that of KWOC, except that for each appearance of a keyword (or descriptor) in URS only the descriptors for the item are printed alongside and not the citation or the annotation. A section of the URS printout is reproduced in Figure 7.

Figure 7

The section of the printout chosen for reproduction in Figure 7 contains the items tagged with COMPUT/IR as a descriptor. The CODEX contains twelve entries tagged with COMPUT/IR. Of these, two contain COMPUT/IR as a "Critical" Descriptor, indicating that "mechanical and electronic developments for information retrieval" is one of the most important facets of the work. Assuming

the CODEX user is interested in only these two items, he would note their identification numbers--AO182 and AO610--and obtain complete citations and accompanying annotations from the CODEX Catalog (Figure 6).

Because the user must locate the keyword and the complete citation in two stages, the URS CODEX is a "double-entry" index-- despite its use of the KWOC format. ¹he computer program could have been written so that the complete citation (and annotation, if desired) would print for every occurrence of a descriptor, but this would have resulted in an enormous publication. As it is, International Affairs requires 1,205 double-column letter-size pages to print the computer output. Only about 200 of these pages are required for the 3,030 item Catalog; the other 1,000 pages contain almost 40,000 alphabetized descriptors and related sets of descriptors. If the descriptors had not been truncated but had been spelled out, de Grazia estimated that "the size of the CODEX would be at least tripled."²⁴

Without question the great value of the International Affairs CODEX lies in the thorough indexing performed on the input. At the same time, however, the in-depth indexing presents problems in using the volume. The basic difficulty is that numerous descriptors produce many index entries and enormous output. In order to hold down the size of the volume, descriptors were truncated and citations were suppressed from being printed with every entry. But these are more or less minor impediments to the use of the CODEX; one soon deciphers the truncations and one can pay the cost of double-entry. The more important problem with the CODEX is that deep indexing causes it to retrieve too much under certain topics.

Many of the URS descriptors produce literally hundreds of entries in the International Affairs CODEX. In the most extreme case, the user who is interested in references dealing with "attitudes (public and private), opinions, or ideology" is confronted with a list of 1,539 references under the Standard Descriptor, "ATTIT." This list encompasses more than half of the items in the entire bibliography and extends for almost forty double-columned pages. Clearly, this index presents a problem to the person interested in attitudes and opinions in international affairs. The same can be said for those interested in "power" (1,235 entries); "education, propaganda, and persuasion" (1,091 entries); and "international

organizations" (1,078 entries). A frequency distribution for descriptors that have one hundred entries or more is given in Table 1.

TABLE 1: Frequency Distribution for URS Descriptors
Retrieved One Hundred Times or More

Number of times retrieved	Number of Descriptors
1,000 or more times	4
750 to 999 times	7
500 to 749 times	15
250 to 499 times	26
100 to 249 times	46
Total	98

As Table 1 shows, ninety-eight descriptors retrieve more than one hundred entries each. In view of the fact that there are only two hundred and sixty-two descriptors in the International Affairs CODEX, more than one-third of them promise little help to the time-pressed researcher who wants to retrieve at most a handful of important references bearing directly on his interests.

This problem of "over-retrieval" of certain topics is not a damning one, and ways are being devised to cope with it in forthcoming volumes in the URS Series. One possible solution is to have "sub-indexes" within the main index. Instead of listing 1,539 items under "ATTIT" without differentiating among them, "ATTIT" could serve as a separate heading for an alphabetical listing of the remaining descriptors in these items. The use of such sub-indexes would add appreciable to the length of the product, but other economies might be introduced to offset the increase. At one extreme, the full set of descriptors could be omitted with only the Catalog numbers given in the alphabetized descriptor list.

Despite problems in its use, de Grazia's Universal Reference System stands as perhaps the most comprehensive index to inter-disciplinary literature on international affairs. International Affairs and the forthcoming volumes in the series will put an important research tool in the hands of researchers who do

not have direct access to the machine-readable input from which the volumes were prepared. But for those who do have direct access to computing facilities, it is becoming clear that published indexes to bibliographical literature are inadequate substitutes for computerized services that can retrieve highly relevant literature tailored to individual needs.

PART III: DOCUMENT RETRIEVAL SERVICES

Alphabetized indexes to literature are intended to be comprehensive, general-purpose publications for selected audiences. In the case of the Cumulative Index to the American Political Science Review, the audience is presumed to be political scientists and other scholars with interests in political science. Scholars interested in the politics of Africa and the Middle East constitute the audience for the KWOC index to that literature. And International Affairs covers literature of interest to students of international relations.

Each of these publications is designed to index its contents from "A" to "Z". This feature makes each a "general-purpose" research tool for anyone in its audience. It is presumed that each publication will be of some use to many different individuals. Because of difficulties in determining in advance which topics will be consulted by its users, a general-purpose index indiscriminantly indexes everything. Because any particular user will be interested in only a small fraction of the entries in a general-purpose index, it often is an inefficient device for serving specific research needs, which are better served by searching and retrieving bibliographical references on an ad hoc basis. This suggests the value in developing continuing services for document retrieval rather than published indexes to entire bibliographies.

Automated document retrieval services have several advantages over bibliographic indexes. One of the obvious ones is in the size of the product. The Cumulative Index to the American Political Science Review has over 10,000 entries; the KWOC index to Africa and the Middle East has almost 4,000; and International Affairs, the length of the index became an important consideration in determining how the literature would be indexed and how much information about each reference would be provided to the user. Because a document retrieval system aims only at providing the user with a manageable number of highly relevant references, its output is small and susceptible to printing in a form

that maximizes convenience to the user.

Another important advantage of a document retrieval service is its ability to produce highly relevant references through more complex searching techniques. To prepare a keyword index, the computer is instructed to search each item for the existence of individual keywords and to prepare an alphabetized list of keywords turned up in its search. However, a computer can also be programmed to search not only for the existence of individual keywords but for certain combinations of keywords: "attitudes" and "aggression", for example. While the number of possible combinations of hundreds of keywords is not infinite, it is very large indeed. To say the least, it would be prohibitively expensive to prepare and publish and index to all these combinations. But searching a bibliography of items to retrieve only those containing both "attitudes" and "aggression" is a relatively simple operation for a computerized document retrieval service.

The great disadvantage of an automated document retrieval system is that it requires ready access to computing equipment.²⁵ This presents a special limitation for researchers at smaller institutions that lack computing facilities. For these individuals, the published bibliographical index is the only real alternative. But even for researchers at institutions which have computing facilities, a document retrieval service is impractical without ready access to the computer. One usually wants to know today what literature bears on his research; he does not want to wait until next week, or even until tomorrow. (In fact, he may want to know immediately, and even this may be possible with the advent of time-sharing computing facilities and remote input-output terminals installed in the researcher's office.)

Major advances remain to be made in developing computerized document retrieval services.²⁶ Nevertheless, some quite sophisticated services are already in existence, such as Harvard's SMART system.²⁷ Most existing services, however, do not deal specifically with references on international relations and comparative politics. Two document retrieval services intended to serve needs in these areas are now under development at Northwestern University. One service aims at retrieving references for the University of Pennsylvania's "International Studies of Values in Politics". The other service proposes the "selective dissemination of information" to members of Northwestern's Council for Inter-Societal Studies.

Both of these services will be discussed in turn.

TRIAL: Technique to Retrieve Information from Abstracts of Literature

A document retrieval service that searches for combinations of related key-words has been proposed for the "International Studies of Values in Politics" (ISVP) project. ISVP involves a program of coordinated field research in four countries: India, Poland, Yugoslavia, and the United States. The project, which began in 1964, is directed by Philip Jacob of the University of Pennsylvania. Philip Jacob and Henry Teune, the project's associate director, had informed me of their need to manage hundreds of bibliographical items that would be assembled by and for scores of research workers on the project. After consultation in the fall of 1964 and the spring of 1965, we devised a document retrieval service that could be tested with an information retrieval program already in use at Northwestern University.²⁸ If the results seemed satisfactory from searching some sample items at Northwestern the program could be adapted to Pennsylvania's computer so the service would be available where it was needed.

Early in the summer of 1965, forty sample bibliography items were sent to Northwestern for keypunching and computer processing. Every item was tagged with several descriptors in the manner of the URS input. But the ISVP set of descriptors was less structured and used whole words instead of truncated terms. When keypunched in the appropriate format for computer processing, the entries looked like the one in Figure 8.

Figure 8

The printout in Figure 8 reveals a punchcard format very much like that used for KWIC and KWOC input. "Authors," "titles," and "sources," are again recorded on different "classes" of cards, indicated in column 75 as classes "1," "2," and "3" respectively. Another class of cards, class "4," is used for the ISVP descriptors. All the cards for the sample items were read by the computer and written on a magnetic tape which constituted the "master file" of bibliography items for the document retrieval service.

The computer program used to search the master file and retrieve specific references is called "TRIAL," an acronym for "Technique to Retrieve Information from Abstracts of Literature."²⁹ The TRIAL program was originally devised for retrieving propositions abstracted from literature on political participation,³⁰

but it since has developed into a far more general information retrieval program. The important feature of the TRIAL program for the ISVP project is its ability to search the master file for combinations of related keywords.

A search of the tape containing the master file is initiated with a program control card that contains a brief description of the search. This description serves only to label the output, and any information may be punched into the control card for this purpose. A sample search for items on the attitudes of elites toward foreign affairs might be labeled, "SEARCH FOR ELITE ATTITUDES ON FOREIGN POLICY." A comma is used to terminate the search label and to alert the computer for the designation of the card "classes" which the user wants to search for keywords or descriptors. For the ISVP test, the computer was instructed to search both class 2 cards (titles) and class 4 cards (descriptors).

The search command itself is communicated to the computer by specifying within parentheses certain keywords and logical connections that must exist between the keywords before an item qualifies for retrieval. After reading the numbers of the card classes to be searched, the machine senses that it is receiving a search command when it reads a left parenthesis. Because any given command may contain more than one combination of keywords arranged in a "nest" of parentheses, the machine continues to read until the number of right parentheses equals the number of left ones. It will then evaluate the logic of the command, working outward from the innermost set of parentheses.

The user of the service must specify the keywords he wants the computer to use in its search. In the sample search mentioned above, the words "elite," "attitudes," "opinions," "foreign policy," and "foreign affairs" might be involved. These terms can be incorporated in a TRIAL search by means of special word commands and logical operators. Simply enclosing a word within parentheses will involve that word in the search. Enclosing any word between slashes (e.g., "/ELITE/") will define the first eight characters of that word as a "root word" and will cause the machine to retrieve any word containing the same combination of characters. The command "/ELITE/" will retrieve "elites" as well as "elite." Joining any pair of words with an asterisk (e.g., "FOREIGN*POLICY") defines the words as a phrase and causes the machine to search for exactly the same phrase.

The real power of the search command lies in the use of the standard logical

operators expressed on the punchcard as follows: .NOT., .OR., and .AND.

Perhaps the use of these operators can best be understood by consulting the sample command in Figure 9, which was constructed for the "elite and foreign affairs" example. The dollar sign ending the command signals the computer to start looking for the next one.

Figure 9

The command in Figure 9 will cause the computer to select only those items tagged with the following descriptors: (1) "elite" or "elites," and (2) "attitudes" or "opinions," and (3) "foreign policy" or "foreign affairs." Only one of the forty items in the ISVP test possessed this combination of descriptors, and it was printed out by the computer as in Figure 10, with the relevant descriptors underlined by hand.

Figure 10

The TRIAL program was used to demonstrate the feasibility and utility of a computerized document retrieval service. On the basis of the test runs made on sample data at Northwestern the ISVP project is moving ahead with plans to operate a similar document retrieval service at the University of Pennsylvania.³¹ At present, the TRIAL program is being rewritten for the computing facilities at Pennsylvania, so researchers on the ISVP project will have ready access to this research tool.

SDI: Selective Dissemination of Information

The bibliographical retrieval service proposed for the International Studies of Values in Politics stands ready to search for specific references on an ad hoc basis; searches are initiated at the request of individual researchers seeking answers to specific questions. It is a logical extension of the service to notify individual researchers automatically of new publications relevant to their current interests by periodically matching abstracts of recent literature against personal sets of keywords. Such a service would keep its users abreast of recent and current publications without requiring frequent trips to the library. The IBM-developed idea of "selective dissemination of information" (SDI) proposes just such a service.³²

In brief, SDI operates as follows. Individual users personally prepare lists

of keywords describing their current interests. Each user's list, which constitutes his "interest profile," is stored on magnetic tape along with interest profiles of others. As new publications appear, they are abstracted and keypunched for storage on computer tape. Periodically, the information recorded on the tape is matched against the interest profiles in the system. When terms in an abstract match terms in a user's interest profile, the computer prints the name of the scholar and the abstract of the document. The abstracts retrieved after each run of the SDI program are sent to the appropriate individuals by mail, notifying them of recent literature presumed to be within their interests. Quite distinct from a general-purpose index to recent publications, such a document retrieval service provides for "selective dissemination of information."

An experimental SDI system has been established at Northwestern University to serve the needs of faculty members engaged in cross-national research. Sponsored by funds from Northwestern's Council for Inter-Societal Studies,³³ the system at present has some 144 articles published mostly during 1966 from eighteen different journals in anthropology, economics, history, psychology, political science, and sociology. A list of the journals, which were chosen for their coverage of topics relevant to inter-societal studies, shows the following currently in the system:

American Anthropologist

American Behavioral Scientist

American Historical Review

American Journal of Economics and Sociology

American Journal of Sociology

American Political Science Review

American Sociological Review

British Journal of Sociology

Comparative Studies in Society and History

Current Anthropology

Economic Development and Cultural Change

Ethnology

Human Organization

International Social Science Journal

Journal of Conflict Resolution

Journal of Personality and Social Psychology

Public Opinion Quarterly

World Politics

The SDI service at Northwestern is still in an early stage of development. Our plans are to increase the journal coverage to meet demands of SDI participants, who presently include only the eighteen members of the Inter-Societal Council, which represents departments and schools within the University that participate in inter-societal studies. Because the Inter-Societal Sdi service has not yet developed into a regular service for participants, the Northwestern experience to date will not be considered further in this paper. Instead, the operation of SDI will be illustrated with reference to searches made against interest profiles of participants in the Computers and Public Policy Making Community Institute.

On March 18, letters were mailed to participants in the Institute asking for information which would help in preparing this paper, scheduled for presentation on April 4, less than three weeks later. The one-page letter briefly described the operation of SDI and stated:

I would like to illustrate the operation of the system by running your interest 'profile' against the articles stored on tape and bringing the computer output to the Institute for distribution and discussion.

Would you please cooperate by providing me with a list of ten terms that describe your current interests? Sample terms for your group might be words like "foreign policy," "aggression," "NATO" "conflict," etc. Combinations of words, like "foreign policy," will be counted as one word. In preparing these terms, it is usually helpful to focus only on the most relevant and rewarding literature.

I have enclosed a pre-addressed postcard to facilitate your response. I respectfully request that you fill out the card and return it to me as soon as possible so that I have enough time to make the computer run before leaving for the Institute.

By Monday, March 28, thirty-seven of forty-five persons in the original mailing had submitted a list of keywords for an SDI search. Mr. Gary Rader reworked these keywords into forms suitable for searching and made the runs on the computer. The search results will be presented and discussed after reviewing some fundamental features of Northwestern's SDI system.

The computer program used in the SDI system is TRIAL, the same program used for the ISVP document retrieval service. In revising his IBM 709 TRIAL

program to operate on Northwestern's new Control Data Corporation 3400 computer. Mr. William Tetzlaff included a new feature in the search logic for SDI applications. In addition to retrieving items according to logical combinations of keywords, a procedure was introduced for "weighting" keywords so that retrieval would occur only if the sum of the weighted keywords in a given abstract equaled or exceeded a previously stated value. This value, or "hit level," is presently set at 100, which causes retrieval when encountering one keyword weighted "100", two weighted "50" each, and so on.

The input to Northwestern's SDI system includes the standard information about author, titles, and sources, which is keypunched in compatible KWIC-KWOC-TRIAL format. In addition, the author's institutional affiliation, when provided, is keypunched on the "author" cards, class "1". The citations and the authors' institutional affiliations are accompanied, when available, by abstracts published with the article, which are keypunched on class "4" cards. When abstracts are not available, which is the case in twelve out of the eighteen journals, keypunchers are given instructions to use the first and last paragraphs, keypunching the first paragraph on class "5" cards and the last paragraph on class "6" cards. (The second and next-to-last paragraphs are substituted if the first or last paragraphs are clearly not suitable as introductions or summaries.) Finally, in the case of review articles, the titles of books reviewed are punched on class "7" cards. A sample of punchcard input of an article from Economic Development and Cultural Change is given in Figure 11.

Figure 11

Keypunching the first and last paragraphs of each article when abstracts are not available is intended to skirt the time-consuming, resource-draining, and generally difficult task of abstracting articles for input to the system. Admittedly, this is a compromise solution to the problem of describing the article for effective retrieval. Some evidence indicates, however, that users find the first and last paragraphs of a article to be more helpful than a professional abstract in determining its value for their research.³⁴ No doubt some relevant articles escape retrieval because the opening and closing paragraphs are used in lieu of abstracts, but extensive abstracting is too costly to consider at this stage of the project.

Using the input as prepared for all 144 articles in the system, what were the results of the SDI search for Institute participants? The master tape of abstracts was searched for each of the thirty-seven institute participants whose keywords were received before March 28. Mr. Rader incorporated their reworked keywords into SDI search instructions for the TRIAL program. To illustrate an actual command which uses most of the features in the program, the instructions prepared for Professor John F. Gange of the University of Oregon have been selected for reproduction in Figure 12.

Figure 12

The label for the search instruction in Figure 12 is "PROF. JOHN F. GANGE--UNIVERSITY OF OREGON". The computer is told to search card classes 2, 4, 5, 6, and 7. The left parenthesis following "7" begins the search command proper. Because the TRIAL program ignores blank spaces in the search command, the keywords can be listed as in Figure 12 to improve readability for humans. Note that the last term -- "SELF*DEFENSE -50," -- is followed by a right parenthesis, which ends the command, and a dollar sign, which starts the computer looking for the next command.

Professor Gange's interest profile in Figure 12 retrieved sixteen abstracts from eight different journals. One of the abstracts retrieved in his search is given in Figure 13. The keywords that caused the abstract to be retrieved are underlined by hand.

Figure 13

Table 2 summarizes the results of the SDI searches for all thirty-seven participants. Out of 144 articles in the system, an average of eight abstracts were retrieved per person. It is for the individual participant to judge whether his "selectively disseminated" abstracts are of any interest to him. In a continuing SDI service, of course, the user can revise his interest profile at will to change the character (and amount) of information he is sent. In this way, a user should be able to learn how to use the service to improve its usefulness in his research.

ADELMAN, IRMA
JOHNS HOPKINS UNIVERSITY
MORRIS, CYNTHIA TAFT
AMERICAN UNIVERSITY

EDCC
EDCC
EDCC
EDCC

A QUANTITATIVE STUDY OF SOCIAL AND POLITICAL DETERMINANTS
OF FERTILITY

EDCC
EDCC

ECONOMIC DEVELOPMENT AND CULTURAL CHANGE
14 (JANUARY, 1966), 129-157

EDCC
EDCC

THE PRESENCE OF A SIGNIFICANT RELATIONSHIP BETWEEN FERTILITY RATES, AND SOCIO-POLITICAL ENVIRONMENT IS WELL KNOWN., HOWEVER, PRECISE INDICATIONS OF THE MAGNITUDES INVOLVED HAVE BEEN LACKING. IN THIS PAPER, AN ATTEMPT IS MADE TO GAIN SOME SEMI-QUANTITATIVE INSIGHTS INTO THE INFLUENCE UPON FERTILITY OF VARIOUS TYPES OF SOCIAL AND POLITICAL CHANGE. PURELY DEMOGRAPHIC VARIABLES HAVE BEEN OMITTED FROM THE ANALYSIS IN ORDER TO TEST THE RELATIVE QUANTITATIVE IMPORTANCE OF SELECTED ASPECTS OF THE TRANSFORMATION OF INSTITUTIONS AND VALUES ASSOCIATED WITH ECONOMIC DEVELOPMENT IN GENERATING ATTITUDES FAVORABLE TO FAMILY LIMITATION.

EDCC
EDCC
EDCC
EDCC
EDCC
EDCC
EDCC
EDCC
EDCC
EDCC
EDCC

IN THIS PAPER, SOME QUANTITATIVE INSIGHTS WERE OBTAINED INTO THE NATURE OF THE RELATIONSHIP BETWEEN CRUDE BIRTH RATES AND TWO ASPECTS OF SOCIAL AND POLITICAL CHANGE COMMONLY ASSOCIATED WITH ECONOMIC DEVELOPMENT-- THE SOCIO-CULTURAL CONCOMITANTS OF THE INDUSTRIALIZATION-URBANIZATION PROCESS (FACTOR I), AND WESTERNIZATION OF POLITICAL INSTITUTIONS (FACTOR II). THE RESULTS OF THE FACTOR ANALYSIS INDICATE THAT APPROXIMATELY 47 PERCENT OF INTER-COUNTRY VARIATIONS IN FERTILITY ARE ASSOCIATED WITH DIFFERENCES IN THE DEGREE OF SPECIALIZATION AND INTEGRATION OF SOCIAL STRUCTURE. AS MIGHT HAVE BEEN EXPECTED, THERE IS A SYSTEMATIC TENDENCY FOR BIRTH RATES TO DECLINE WITH THOSE SOCIAL CHANGES WHICH CHARACTERISTICALLY ACCOMPANY INDUSTRIALIZATION. DIFFERENCES IN THE STRENGTH OF POLITICAL WESTERNIZATION (FACTOR II) ACCOUNT FOR AN ADDITIONAL 11 PERCENT OF CROSS-COUNTRY VARIATIONS IN BIRTH RATES. IN GENERAL, COUNTRIES CHARACTERIZED BY LOWER NATALITY TEND TO BE THOSE IN WHICH POLITICAL SYSTEMS OF THE WESTERN TYPE ARE MORE HIGHLY EVOLVED. ON THE OTHER HAND, THE INCLUSION OF A BROAD MEASURE OF ECONOMIC STRUCTURE IN THE ANALYSIS DOES NOT ADD SIGNIFICANTLY TO THE EXPLANATION OF FERTILITY. (SECOND FROM LAST PARAGRAPH)

[illegible]

TABLE 2: List of Institute Participants and the Number
of Abstracts Retrieved for Each

Name of Participant	Number of Abstracts
Edwin Adams	4
Howard Ball	0
Abraham Bargman	12
Davis Bobrow	6
Susan Bobrow	10
Paul Borel	5
Harold Borko	0
Keith Clark	2
Walter C. Clemens, Jr.	3
William D. Coplin	7
John F. Gange	16
Ralph M. Goldman	9
Morton Gorden	20
Alfred James Hartzler	2
Samuel P. Hayes	11
Peter Heintz	19
Ole Holsti	9
James G. Hoofnagle	12
Fisher Howe	3
Julius Katz	1
James A. McDevitt	5
Knight McMahan	3
Jeffrey Milstein	4
Bernard S. Morris	4
Eric C. Orme	5
David Osborn	7
Dallas A. Pilliod	4
Ithiel de Sola Pool	9
James E. Price	1
Peter H. Rohn	0
Richard C. Rowson	5
Rudolph Rummel	22
Robert Strausz-Hupe	17
Leo Sweeney	13
William J. Trainor, Jr.	28
John Vance	4
William Welsh	12

CONCLUSION

Keyword indexing of bibliographical items by computer is an established document retrieval technique. It is fast, inexpensive, and easily executed. If titles of interest are relatively descriptive or if they can be made more descriptive after some editing, then keyword indexing by computer offers a practical means for producing useful general-purpose bibliographical indexes at low cost. At considerably more expense, each item in the bibliography can be tagged with a number of descriptor terms to aid retrieval through more complete indexing. But a general-purpose index produced from a large set of items tagged with many descriptors constitutes a bulky volume - as demonstrated by International Affairs, Volume I of the Universal Reference System. Despite discouragingly long lists of index entries, such comprehensive indexes serve worthy purposes when published, and they are especially useful to researchers who do not have access to computing equipment. If computers are available, however, they can be incorporated into document retrieval services which can conduct more powerful searches and thereby retrieve less chaff and more wheat from large files of bibliographical items.

FOOTNOTES

¹ Those interested in the technology of information retrieval may wish to consult these recent texts: Charles P. Bourne, METHODS OF INFORMATION HANDLING (New York: Wiley, 1963); Allen Kent, TEXTBOOK ON MECHANIZED INFORMATION RETRIEVAL (New York: Wiley, 1962); Joseph P. Spiegel and Donald Walker (eds.) INFORMATION SYSTEM SCIENCES (Washington: Spartan Books, 1965).

² At least three non-computer methods of document retrieval deserve to be mentioned: (1) the familiar "keysort" cards of Royal McBee Corporation; (2) the "Termatrex" system of Jonker Business Machines, Inc. (Gaithersburg, Maryland); and (3) various microfilm or "microfiche" cameras and readers, but in particular the "MIRACODE" system of Recordak, a division of Eastman Kodak, Co. Information about all of these are available from the manufacturers. For a more specialized review of MIRACODE's application to the problem of information retrieval, see Kenneth Janda, "MIRACODE: A 16 mm. Microfilm Information Retrieval System," paper prepared for delivery at the Third Technical Conference of the Council of Social Science Data Archives, University of Michigan, May 10-12, 1966 (scheduled to be published in the conference proceedings).

³ See, for example, the index to my Data Processing: Applications to Political Research (Evanston: Northwestern University Press, 1965), which was prepared on a computer with a keyword indexing program. For a different approach, see Susan Artandi, "Automatic Book Indexing by Computer," American DOCUMENTATION, 15 (October, 1964), 250-257

⁴ One such program for the popular IBM 1401 computing system is described in "Keyword-in-Context (KWIC) Indexing Program for the IBM 1401 Data Processing System," Reference Code 1401-CR-02X (White Plains, N. Y.: International Business Machines, 1963). A recent revision of earlier KWIC programs for computers in the IBM 700 series had been prepared by Professor James S. Angaard of Northwestern University's Department of Electrical Engineering. Angaard's program, called "EIKWIC," is available through SHARE.

⁵ The discussion of keyword indexing is reprinted with revisions from Kenneth Janda, "Keyword Indexing for the Behavioral Sciences," THE AMERICAN BEHAVIORAL SCIENTIST, 7 (June, 1964), 55-58.

⁶ H. P. Luhn is generally considered to be the father of keyword indexing. See his "Keyword-in-Context Index for Technical Literature (KWIC Index)," IBM Advanced Systems Development Division Report RC-127, August, 1959.

⁷ Other social science applications of keyword indexing include Murray A. Straus and Susanne C. Graham, CUMULATIVE INDEX TO MARRIAGE AND FAMILY LIVING, Volume 1-24 (1939-62); CURRENT STATE LEGISLATION INDEX (American Bar Foundation, Indianapolis: Bobbs-Merrill Company, beginning publication in 1962); and the comprehensive index to publications in the field of social welfare from 1924 to 1962, prepared by Joe R. Hoffer,

FOOTNOTES (continued)

Executive Secretary of the National Conference on Social Welfare, Columbus, Ohio.

⁸ Kenneth Janda (ed.), CUMULATIVE INDEX TO THE AMERICAN POLITICAL SCIENCE REVIEW, VOLUMES 1-57: 1906-1963 (Evanston: Northwestern University Press, 1964).

⁹ There is no practical limit to the number of authors, length of titles, or amount of publication information that can be recorded on punchcards for input to the computer. Additional cards within a given "author," "title," or "source" classification can be used as needed. The standard format for recording bibliographic information on punchcards for keyword indexing is given in the IBM General Information Manual E20-8091, "Keyword-In-Context (KWIC) Indexing."

¹⁰ The particular program used to process this information was written by James S. Aagaard. See footnote 4.

¹¹ Aagaard's computer program performed the alphabetizing on the basis of the first seventeen characters following the first letter of the keyword, which accounts for alphabetizing the name keywords according to the next word in the titles.

¹² For a comparison of the two methods, see B.B. Lane, "Key Words in-- and out of--Context," SPECIAL LIBRARIES, (January, 1964), 45-46.

¹³ This bibliography was prepared for my research on comparative political parties. See Kenneth Janda, "Retrieving Information for a Comparative Study of Political Parties," in William J. Crotty (ed.), APPROACHES TO THE STUDY OF PARTY ORGANIZATION (Chicago: Rand McNally, forthcoming).

¹⁴ The KWOC index was processed on Professor James S. Aagaard's new "Bibliographic Data Processor" program for the IBM 709, which is being rewritten for the Control Data Corporation 3400 computer. Aagaard's program contains a number of options which can be exercised simply by calling for them with the appropriate control cards. KWIC and KWOC processing are two of the options; KWAR (keywords without reference code) and CORK (combinations of related keywords) are two others. In addition, the program allows great flexibility in determining which classes of cards shall be searched for keywords and which classes of cards shall be printed in the output.

¹⁵ KWOC's ability to print the entire title regardless of length lends this form of indexing to applications outside of bibliographic indexing. At Northwestern, KWOC indexes have been prepared to all roll call votes in the 15th Session of the United Nations General Assembly. As an undergraduate major in political science at Northwestern, Mr. David Hodge prepared KWOC indexes to all contested roll call votes in the Kansas Senate from 1957 to 1965.

FOOTNOTES (continued)

- ¹⁶ Lane, op cit.
- ¹⁷ Donald H. Kraft, "A Comparison of Keyword-In-Context (KWIC) Indexing of Titles with a Subject Heading Classification System," AMERICAN DOCUMENTATION, 15 (January, 1964), 48-52. Kraft's analysis of 3,248 entries in an issue of THE INDEX TO LEGAL PERIODICALS and an issue of THE INDEX TO LEGAL THESIS AND RESEARCH PROJECTS disclosed that 89.57. had titles Kraft judged suitable for keyword indexing.
- ¹⁸ The bibliographic items, published in a section called "New Studies: A Guide to Recent Publications in the Social and Behavioral Science, " have been selected and annotated by the journal staff in a periodic search of hundreds of journals and reviews, including many published outside of the United States.
- ¹⁹ International Affairs is published by the Universal Reference System, a division of Metron, Inc.; 80 East Eleventh Street; New York, N.Y. 10003.
- ²⁰ Ibid., page ix.
- ²¹ Ibid., page xi.
- ²² Ibid., pages xv-xix.
- ²³ Ibid., page vii. The programming was done by Datistics, Inc. of Jersey City, N.J.
- ²⁴ Ibid., page xii.
- ²⁵ Cost of computer time can also be a major disadvantage of computerized document retrieval system, but cost of computer time and availability of free time range widely from installation to installation.
- ²⁶ A major advantage in computing "hardware"--the machinery itself--would be the development of inexpensive and effective optical scanning equipment for translating printed material directly into punchcard form. A major advance in computing "software"--programming systems-- would be the development of fast and accurate programs for translating foreign languages into English.
- ²⁷ G. Slaton and M.E. Lesk, "The SMART Automatic Document Retrieval System-- An Illustration," COMMUNICATIONS OF THE ACM, 8 (June, 1965), 391-398.
- ²⁸ The program in question was written by Mr. William H. Tetzlaff of the Computing Center at Northwestern University. See his "TRIAL: Technique for Retrieving Information from Abstracts of Literature--A Program for the IBM 709/90/94," paper delivered at the SHARE XXVI Meeting, Applied Management Sciences Project Committee; San Diego, California; February 28 - March 4, 1966.

FOOTNOTES (continued)

²⁹ The program is discussed in depth with particular reference to literature on political participation in Kenneth Janda and William H. Tatzlaff, "TRIAL: A Computer Technique for Retrieving Information from Abstracts of Literature," BEHAVIORAL SCIENCE, 11 (November, 1966), forthcoming.

³⁰ See Lester W. Milbrath and Kenneth Janda, "Computer Applications to Abstraction, Storage, and Recovery of Propositions from Political Science Literature," paper delivered at the 1964 Annual Meeting of the American Political Science Association, Chicago.

³¹ Some material on the TRIAL Program that has appeared elsewhere was reprinted in the "Working Papers" prepared by the American collaborators for use at the First International Round Table of the International Studies of Values in Politics: Dubrovnik, Yugoslavia; June 22 - July 3, 1965.

³² C. B. Hensly, "Selective Dissemination of Information (SDI): State-of-the-Art in May, 1963." Copies available from the author. Address: IBM, 2651 Strang Boulevard: Yorktown Heights, New York.

³³ I wish to thank the Council on Inter-Societal Studies and its director Professor Richard D. Schwartz, for supporting the project through its experimental stage.

³⁴ Allen Kent, "The Information Retrieval Game," Paper delivered at a Conference on Electronic Information Handling; Pittsburgh, Pennsylvania; October 7-9, 1964.

Information Retrieval: A Proposal

1. The attached paper by Prof. Kenneth Janda does the following:
 - (a) reviews types of information retrieval systems;
 - (b) evaluates the systems;
 - (c) describes in some detail the applications of the system to international and comparative politics.
2. Of particular interest is his description of the retrieval system which was employed by the International Studies of Values in Politics. The work, which was done in the spring and summer of 1965, is described in detail on pp. 18-21.
3. Although effort and thought have gone into a system for the ISVP project, lack of personnel has delayed implementation. It was envisaged that retrieval and dissemination would involve the following procedures:
 - (a) development of a set of key descriptors for the project as a whole;
 - (b) collection of relevant materials by each center;
 - (c) coding of the materials by each center according to the selected key words;
 - (d) a bibliographic print-out of materials at set intervals.
4. At present the program for Selective Dissemination of Information (SDI) has been financed by the Inter-societal program at Northwestern University and organized by Professor Janda. This means that about 22 journals in the social sciences will be coded according to titles. Individual researchers and organizations will give to the system an "interest profile". The "interest profile" is a set of words. On the basis of this set of words these 22 journals will be periodically searched and the materials meeting the "interest profile" will be printed out in any number of copies.

5. Professor Janda is prepared to accept an interest profile for the ISVP program. The cost for each search will be about \$19, depending on the complexity of the interest profile and not upon the number of items located.
6. The proposal then is to construct a set of words which will comprise an interest profile for the project. Certain words such as "values", "decision-making", and "integration" will, of course, be included. Others should be added. It will be possible to know that all centers are collectively informed about the relevant current literature contained in 22 journals on a systematic, regular basis.
7. This simple procedure does not meet all of the goals which retrieval systems could provide. It does function for current literature. It is now feasible, cheap and ready to go on 1966 journals.
8. In other papers Prof. Janda has outlined on other methods of information retrieval. Among them is a system whereby each paragraph or page is coded on a six digit basis. The code then, in a matter of seconds, searches up to 2000 pages and locates relevant material. The material can be immediately read and, if desired, a copy of the page can be made. Parts of the Human Relations Area File are now recorded in this system.
9. It is proposed that we take advantage of the simple SDI system. Information retrieval is in a revolutionary stage and now is producing tangible benefits.