

AD 653726

**TECHNICAL WRITING:
A KEY TO COMPUTERIZED
INFORMATION RETRIEVAL**

PREPARED BY
SO JONES
CORPORATE LIBRARIAN
DOUGLAS AIRCRAFT COMPANY, INC.

PRESENTED TO
THE 14TH INTERNATIONAL TECHNICAL COMMUNICATIONS
CONFERENCE OF THE SOCIETY OF
TECHNICAL WRITERS AND PUBLISHERS
CHICAGO, ILLINOIS
24 - 27 MAY 1967

 JUN 28 1967

DOUGLAS CORPORATE OFFICE

ARCHIVE COPY

F
24

Best Available Copy

TECHNICAL WRITING:
A KEY TO COMPUTERIZED INFORMATION RETRIEVAL

by

Shirli O. Jones
Corporate Librarian
Douglas Aircraft Company, Inc.

ABSTRACT

The future success of mechanized retrieval of information from published material lies, in a large part, in the hands of the technical writer. The context is only as usable as its format when current retrieval methods are used. Examples are given of confusion arising through nondefinitive titles, particularly where permuted indexes, such as Key-Word-In-Context (KWIC) and Key-Word-Out-of-Context (KWOC), are used.

The appearance of an article in several different publications is another source of difficulty for the information retrieval system. How the creation of separate entries for the same item can damage the integrity of an information system is shown.

A plea is made to the technical writers for understanding the needs of the second generation of reader. Steps which can be taken without censoring the language by replacing all synonyms and near-synonyms with a single term are indicated. The use of meaningful titles, standardization of citations, and several other areas are explored from the viewpoint of the information specialist.

FIGURES

1	KWIC and KWOC Indexes	4
2	Title Repetition by Word	5
3	KWOC Index With Terms Repeated	6
4	List of Words Prevented From Indexing	7
5	Effects of Hyphen Usage Upon Computer Sort	8
6	Effects of Singular-Plural Usage Upon Computer Sort	9
7	Index Card With Multiple Corporate Authors	12
8	Variations of Single Corporate Author	14
9	Automatic Index	16

Preceding Page Blank

INTRODUCTION

Each year, more and more persons decry the proliferation of the printed word. Available statistics prove that unless restraint is applied, mankind will be submerged in published information by 1975. Concern over how to obtain essential information from this vast amount of material results in a rash of papers such as this. Information specialists involved with the retrieval of necessary data bemoan the unnecessary generation of papers, but of more importance is more careful consideration in both the preparation and the publication of all papers. Such a specialist is not in a position to determine the relative value of one document in comparison with another, but he is in a position to help establish guidelines suitable for advanced computer techniques that will provide the keys to necessary storage and retrieval capabilities for the future. Although information retrieval systems covering all degrees of complexity have been introduced, none can produce a product better than the original material. Therefore, authors and technical writers must cooperate with the information specialist to ensure the rapid and effective retrieval of valuable data.

Information retrieval, in general, relies on two forms of index--subject headings and concept coordination. The more familiar form of the subject heading index is the public library card. The concept coordination system uses "uni-terms" or "descriptors"--a single term or concept for every major topic in a document. These terms can then be coordinated with other terms or concepts for flexible, specialized retrieval. Bibliographic information, such as personal author, corporate author, and title, is handled in much the same way, regardless of the type of retrieval system used. The equipment used in creating and manipulating the index can range all the way from a typist typing on 3 x 5 file cards to a fully computerized system for storage and automated retrieval in various formats. In every case, the technical writer's importance is evident, but it is never so obvious as in the computerized storage and retrieval system. Thus, the problems associated with this area require special emphasis.

DISCUSSION

Since 1959, when Douglas actively began to design its mechanized information retrieval program, many different systems have been developed. Government agencies, such as NASA's Scientific & Technical Information Division and DOD's Defense Documentation Center, have built massive systems utilizing the most sophisticated of computer equipment. Several private industries have expended vast quantities of time and money to provide their personnel with retrieval capabilities. Organizations such as the American Society for Metals and the American Chemical Society have called upon their professional talents to produce systems designed to cope with the complexities of their scientific disciplines. These are all major efforts, but the needs of the smaller systems are of equal importance. And either type of system must face much the same problems. Some of these problems, which will be described in detail, are the responsibility of the technical writers and/or publishers.

One of the most common forms of mechanized index is the Key-Word-In-Context or KWIC index, and a variation called KWOC, or Key-Word-Out-of-Context (Figure 1). As can be seen, these indexes consist of a sorting of the title with each word repeated in alphabetical order. The title will appear as many times as there are significant words in the title (Figure 2). The KWOC index provides a more readable format (Figure 3) but follows the same principle. To prevent listings of insignificant terms (such as "the" and "of"), a stop list is input to the computer. Chemical Titles employs a stop list of 1,335 "insignificant" words (Figure 4). Examination of this list indicates that titles such as "Laboratory Systems & Procedures" would be completely eliminated.

Another problem occurs with the inconsistent use of hyphens. In Figure 5, there are several interesting entries. Under "low," the first inconsistency occurs with "low energy"--two entries are unhyphenated, three hyphenated. Next is "low molecular weight"--two entries without hyphens, one with.

IC LESIONS IN THE BROAD	BREADED BRONZE TURKEY/ THE INFLUENC	7310
ISE/	BREATH HOLDING AT BEGINNING OF EXERC	5742
OF EXERCISE AND O ₂ ON	BREATH HOLDING/ INFLUENCE	5740
CREATINE TREATMENT UPON	BREATHING ENERGETIC OUTPUT AND AEROB	4984
KIN OF THE COMMON SOLE/	BREATHING MOVEMENTS IN ENTODELLA-SO	8326
GITAL PRESSURE AND DEEP	BREATHING IN PATIENTS WITH ADAMS STO	5522
EFFECT OF O ₂	BREATHING ON PULMONARY COMPLIANCE/	5745
SYSTEMIC EFFECTS AFTER	BREATHING POTENTIATED MEDICATED AEROSOLS/	5754
URING POSITIVE PRESSURE	BREATHING/ DIAPHRAGM ACTIVITY AND TH	5734
P AS RELATED TO SEASON,	BREED, SEX, AND SEMEN QUALITY/ THE T	7254
THE DISPERSION AND THE	BREEDING BEHAVIOR OF THE RAVEN WITH	8544
LEAST TERN/ THE	BREEDING BIOLOGY AND ETHOLOGY OF THE	4595
GATIONS OF THE CHOICE OF	BREEDING BIOTYPE/ THE STATE OF DEVEL	4593
SE FREE PATHOGEN FREE	BREEDING COLONY/ PROGRESS REPORT. DI	7371
LICATION OF GENETICS TO	BREEDING DOMESTIC ANIMALS/ APP	6566
CAL PROBLEMS IN LUCERNE	BREEDING IN CONNECTION WITH FERTILIT	7908
E HYBRID FORAGE SORGHUM	BREEDING IN HUNGARY/ THE PRESENT POS	7907
FLORIDA/ AUTUMNAL	BREEDING OF BOAT-TAILED GRACKLES IN	8553
TAMIC ACID FORMATION IN	BREVI BACTERIUM-FLAVUM/ SIGNIFICANCE	6697
CAL COMPOSITION OF SAKE	BREWING WATER AND THEIR GROUPING/ ST	5159
OUPIING/ STUDIES ON SAKE	BREWING WATER. CHEMICAL COMPOSITION	5159
WATER/ STUDIES ON SAKE	BREWING WATER. RELATIONSHIP AMONG CO	5160
MONG COMPONENTS OF SAKE	BREWING WATER/ STUDIES ON SAKE BREW	5160
CK BY ELECTRONIC MEANS/	BRIDGING OF INTERRUPTED ATRIO VENTRI	5436
THE CURING PROCESS FOR	BRIGHT-LEAF TOBACCO/ STEADY-STATE TH	7848
E RESPONSES IN A SIMPLE	BRIGHTNESS DISCRIMINATION UNDER DIFF	4525
A DEVICE FOR FEEDING	BRINE SHRIMP TO FISHES/	4685
OF PLANT PHYSIOLOGISTS,	BRISBANE MAY 1961/ LEAF TEMPERATURE	7716
OF PLANT PHYSIOLOGISTS,	BRISBANE MAY 1961/ THE CUTICLE IN EU	7697
OF PLANT PHYSIOLOGISTS,	BRISBANE MAY 1961/ THE EFFECTS OF MO	7703
THE WITH HIGH MOUNTAIN	BRISKET DISEASE AND IN EXPERIMENTAL	7274
THE GENUS ASELLUS IN	BRITAIN/	8362
CIES OF LITHOSID NEW TO	BRITAIN/ PELOSIA-OBTUSA-HERRICH-SCHA	8419
N POTATO CROPS IN GREAT	BRITAIN, 1952-60/ EXTENT OF PROTECTI	8199
MAN-MADE ACTIVITIES IN	BRITISH COLUMBIA/ THE EFFECTS ON FRE	4643
IZATION IN THE HANDS OF	BRITISH FISH FILLETTERS/ COLD VASODIL	6334
IN THE WEST INDIES AND	BRITISH GUIANA/ EASTERN EQUINE ENCEP	7487
BIOLOGICAL FLORA OF THE	BRITISH ISLES. AARHENATHERUM-ELATIUS	7482
BIOLOGICAL FLORA OF THE	BRITISH ISLES. FRAXINUS-EXCELSIOR-L/	7690
ABERRATIONS OF	BRITISH LEPIDOPTERA/	8413
NOIES, THE GUIANAS, AND	BRITISH-HONOURAS/ PRESENT AND POTENT	7850
YTIUS-EDULIS-L. IN THE	BRITISH-ISLES AND THEIR RELATIONSHIP	8448
CLEROTIC LESIONS IN THE	BROAD BREASTED BRONZE TURKEY/ THE IN	7310
OF THE DISTRIBUTION OF	BROAD-LEAVED EVERGREEN TREES, BASED	4544
F ISOLATED FRACTIONS OF	BROKEN CHLOROPLASTS/ HILL ACTIVITY O	7728
RAIN DIFFERENTIATION OF	BROMEGRASS MOSAIC VIRUS/ PURIFICATIO	8157
TWO VARIETIES OF SMOOTH	BROMEGRASS/ EFFECT OF CERTAIN FERTIL	7916
OL STUDIES OF SOME NEED	BROMEGRASSES/ LIFE CYCLES AND CONTR	7958
A PROTEOLYTIC ENZYME,	BROMELAIN/ SYSTEMIC BIO CHEMICAL CHA	6259
SPECIAL REFERENCE TO 5-	BROMO DEOXY URIDINE/ IONIZATION OF D	4284

Figure 2. Title Repetition by Word

NON-IRRADIATED	ABSORPTION OF B-GLUCOSE BY SEGMENTS OF INTEREST FROM ACTIVE AND HIBERNATING IRRADIATED AND NON-IRRADIATED GROUND SQUIRRELS, CITELLUS 141	NUCLEAR	ETIC BLACKOUT FOLLOWING A HIGH ALTITUDE NUCLEAR CRYSTATION
NON-ISOTHERMAL	NON-IRRADIATED GROUND SQUIRRELS, CITELLUS 141	NUCLEAR	ACCURATE NUCLEAR FUEL BURNUP ANALYSIS
NON-LINEAR	ISOTHERMAL CORRELATIONS IN A NON-ISOTHERMAL PLASMA	NUCLEAR	APPLICATION OF NUCLEAR POWER SUPPLIES TO SPACE SYSTEMS
NON-POLY	INVESTIGATION OF MICROSCOPIC NON-LINEAR EFFECTS UTILIZING FERROMAGNETIC MATERIALS	NUCLEAR	CAOLINAS-VIRGINIA NUCLEAR POWER ASSOCIATES, INC., RESEARCH AND DEVELOPMENT PROGRAM QUARTERLY PROGRESS REPORT FOR THE PERIOD APRIL - JUNE 1962
NON-MILITARY	BIBLIOGRAPHY AND TABULATION OF CAMPING PROPERTIES OF NON-METALLIC MATERIALS	NUCLEAR	COMPUTER PROGRAMS FOR OPTIMUM START-UP OF NUCLEAR PROPULSION SYSTEMS
NON-MOVING	NOTES ON NON-MILITARY MEASURES IN CONTROL OF INSURGENCY	NUCLEAR	DOSE-TIME-DISTANCE CURVES FOR CLOSE-IN FALLOUT FROM LOW YIELD LAND-SURFACE NUCLEAR DETONATIONS
NON-RELATIVISTIC	JUDGMENTS OF VISUAL VELOCITY AS A FUNCTION OF THE LENGTH OF OBSERVATION TIME OF MOVING OR NON-MOVING STIMULI	NUCLEAR	EXTENDED CERAMIC NUCLEAR FUEL DEVELOPMENT PROGRAM
NON-SIMILAR	TABLES OF NON-RELATIVISTIC ELECTRON TRAJECTORIES FOR FIELD EMISSION CATHODES	NUCLEAR	FEASIBILITY DETERMINATION OF A NUCLEAR THERMOELECTRIC SPACE POWER PLANT
NONDESTRUCTIVE	NON-SIMILAR NUMERICAL METHODS OF SOLUTION FOR ELECTRODE BOUNDARY LAYERS IN A CROSSED FIELD ACCELERATOR	NUCLEAR	HIGH-ENERGY NUCLEAR PHYSICS RESEARCH PROGRAM
NONDESTRUCTIVE	NONDESTRUCTIVE SYSTEM FOR INSPECTION OF FIBER GLASS-REINFORCED PLASTIC MISSILE CASES	NUCLEAR	HIGH-ENERGY NUCLEAR REACTIONS OF NIOBIUM WITH INCIDENT PROTONS AND HELIUM IONS
NONDISSIPATIVE	X-RAY IMAGE SYSTEM FOR NONDESTRUCTIVE TESTING OF SOLID PROPELLANT MISSILE CASE WALLS AND BELLOMETS	NUCLEAR	INVESTIGATIONS ON THE DIRECT CONVERSION OF NUCLEAR FISSION ENERGY TO ELECTRICAL ENERGY IN A PLASMA DIODE
NONEQUILIBRIUM	MAGNETOHYDRODYNAMIC STABILITY OF VORTEX FLOW - A NONDISSIPATIVE, INCOMPRESSIBLE ANALYSIS	NUCLEAR	NUCLEAR SUPERHEAT DEVELOPMENT PROGRAM
NONLINEAR	SCALE EFFECTS FOR NONEQUILIBRIUM CONVECTIVE HEAT TRANSFER WITH SIMULTANEOUS GAS PHASE AND SURFACE CHEMICAL REACTIONS. APPLICATION TO HYPERSONIC FLIGHT AT HIGH ALTITUDES	NUCLEAR	PRODUCTION OF TRITIUM BY CONTROLLED NUCLEAR EXPLOSIONS IN SALT. I. LABORATORY STUDIES OF ISOTOPIC EXCHANGE OF TRITIUM IN THE HYDROGEN-WATER SYSTEM
NONLINEAR	APPLICATION OF VARIATIONAL EQUATION OF MOTION TO THE NONLINEAR VIBRATION ANALYSIS OF HOMOGENEOUS AND LAYERED PLATES AND SHELLS	NUCLEAR	STRAINING EFFECT OF NUCLEAR EXPLOSION
NONLINEAR	EXTENSIONS IN THE SYNTHESIS OF TIME OPTIMAL OR BANG-BANG NONLINEAR CONTROL SYSTEMS. PART I. THE SYNTHESIS OF QUASI-STATIONARY OPTIMUM NONLINEAR CONTROL SYSTEMS	NUCLEAR	THE NUCLEAR PROPERTIES OF BISMUTH
NONLINEAR	EXTENSIONS IN THE SYNTHESIS OF TIME OPTIMAL OR BANG-BANG NONLINEAR CONTROL SYSTEMS. PART II. THE SYNTHESIS OF QUASI-STATIONARY OPTIMUM NONLINEAR CONTROL SYSTEMS	NUCLEAR	VARIATIONS IN THE TOTAL ELECTRON CONTENT OF THE EARTH'S IONOSPHERE AFTER THE HIGH ALTITUDE NUCLEAR EXPLOSION
NONLINEAR	NONLINEAR FLEXURAL VIBRATIONS OF SANDWICH PLATES	NUMBERS	480A MARITIME NUCLEAR STEAM GENERATOR
NONLINEAR	OPTIMUM NONLINEAR CONTROL FOR ARBITRARY DISTURBANCES	NUMBERS	THE ESTIMATION PROBLEM IN NULL-ZONE RECEIPTION FEEDBACK SYSTEMS
NONRECURRENT	A TECHNIQUE FOR BROAD-BAND TELEMETRY OF WIDE CURRENT PULSES	NUMBERS	FUNDAMENTAL SOLUTION TO DIFFUSION BOUNDARY LAYER EQUATION FOR NEARLY SEPARATED FLOW OVER SOLID SURFACES AT VERY LARGE PRANDTL NUMBERS
NONUNIFORM	ELECTROMAGNETIC SCATTERING FROM A SPHERICAL NONUNIFORM MEDIUM. PART I. GENERAL THEORY	NUMBERS	LOCAL PRESSURE DISTRIBUTION ON A BLUNT DELTA WING FOR ANGLES OF ATTACK UP TO 95-DEGREES AT WAC NUMBERS OF 3.6 AND 4.7
NONUNIFORM	ELECTROMAGNETIC SCATTERING FROM A SPHERICAL NONUNIFORM MEDIUM. PART II. GENERAL THEORY	NUMBERS	A MAINTENANCE PROGRAM FOR MECHANICAL CONTROL SYSTEMS ON MACHINE TOOLS
NORMAL	PROBABILITY INTEGRALS OF MULTIVARIATE NORMAL AND MULTIVARIATE-T	NUMBERS	NUMERICAL SOLUTION OF THE DIRECT PROBLEM OF A PRIORI BOUNDS ON THE SINGULARIZATION ERROR IN THE NUMERICAL SOLUTION OF THE DIRECT PROBLEM
NORMAL	RESONANCE ABSORPTION OF GAMMA-RAYS IN NORMAL AND SUPERCONDUCTING TIN	NUMBERS	NON-SIMILAR NUMERICAL METHODS OF SOLUTION FOR ELECTRODE BOUNDARY LAYERS IN A CROSSED FIELD ACCELERATOR
NORMS	NORMS FOR ARTIFICIAL LIGHTING	NUMBERS	MANIPULATION OF ARGUS AND ITS EFFECTS ON HUMAN VESTIBULAR NYSTAGMUS INDUCED BY CALORIC IRRADIATION AND ANGULAR ACCELERATIONS
NORTH	FACTORS INFLUENCING VASCULAR PLANT ZONATION IN NORTH CAROLINA SALT MARSHES	NUMBERS	A SAFETY REVIEW OF THE OAK RIDGE CRITICAL EXPERIMENT FACILITY
NORTH	SOUND STUDIES OF THE DEEP SCATTERING LAYER IN THE NORTH PACIFIC	NUMBERS	DRAW OF OBJECTS IN PARTICLE-LADEN AIR FLOW FROM MASS IV, BLUNT BODIES AND COMPRESSIBILITY EFFECTS
NORTH	THE DEVELOPMENT OF RESCUE AND SURVIVAL TECHNIQUES IN THE NORTH AMERICAN ARCTIC	NUMBERS	TUNTO FOREST SEISMOLOGICAL OBSERVATORY
NOSE	THE FLORA OF HEALTHY DOGS. I. BACTERIA AND FUNGI OF THE NOSE, THROAT, AND LOWER INTESTINE	NUMBERS	A SAMPLE TEST EXPOSURE TO EXAMINE CORROSION AND FOULING OF EQUIPMENT INSTALLED IN THE OCEAN
NOZZLE	FABRICATION OF PYROLYTIC GRAPHITE ROCKET NOZZLE COMPONENTS	NUMBERS	OCEANOGRAPHIC CRUISE TO THE BEIRING AND CHUMCHIC SEAS, SUMMER 1949. PART I. SEA FLOOR STUDIES
NOZZLE	FABRICATION OF PYROLYTIC GRAPHITE ROCKET NOZZLE COMPONENTS	NUMBERS	OCEANOGRAPHIC AND UNDERWATER ACOUSTICS RESEARCH
NOZZLE	FABRICATION OF PYROLYTIC GRAPHITE ROCKET NOZZLE COMPONENTS	NUMBERS	OCEANOGRAPHIC CRUISE TO THE BEIRING AND CHUMCHIC SEAS, SUMMER 1949. PART IV. PHYSICAL OCEANOGRAPHIC STUDIES. VOL. I. DESCRIPTIVE REPORT
NOZZLE	THIRD SYMPOSIUM ON ADVANCED PROPULSION CONCEPTS SPONSORED BY UNITED STATES AIR FORCE OFFICE OF SCIENTIFIC RESEARCH AND THE GENERAL ELECTRIC COMPANY FLIGHT PROPULSION DIVISION (CONCLUDED), 1961 OCTOBER 2-4, 1962. PLASMA FLOW IN A MAGNETIC ARC NOZZLE	NUMBERS	OCEANOGRAPHIC CRUISE TO THE BEIRING AND CHUMCHIC SEAS, SUMMER 1949. PART I. PHYSICAL OCEANOGRAPHIC STUDIES. VOL. I. DESCRIPTIVE REPORT
NOZZLES	HEAT TRANSFER AND PARTICLE TRAJECTORIES IN SOLID-ROCKET NOZZLES	NUMBERS	OCEANOGRAPHIC CRUISE TO THE BEIRING AND CHUMCHIC SEAS, SUMMER 1949. PART IV. PHYSICAL OCEANOGRAPHIC STUDIES. VOL. 2. DATA REPORT
NOZTC	DEVELOPMENT AND STANDARDIZATION OF FORMS 3 AND 4 OF THE NOZTC CONTRACT STUDENT SELECTION TEST	NUMBERS	OCEANOGRAPHIC CRUISE TO THE BEIRING AND CHUMCHIC SEAS, SUMMER 1949. PART IV. PHYSICAL OCEANOGRAPHIC STUDIES. VOL. 2. DATA REPORT
NOZTC	EVALUATION OF NOZTC AVIATION INDICATORINATION FIELD TOURS FOR 1961-1962	NUMBERS	PROCEEDINGS OF INTERINDUSTRIAL OCEANOGRAPHIC SYMPOSIUM (NO. 1), BURBANK, CALIFORNIA, 5 JUNE 1962
NUCLEAR	A 709B CODE FOR THE CALCULATION OF ELECTRODYNAMIC	NUMBERS	RUBBER ELASTICITY IN HIGHLY CROSSLINKED SYSTEMS

Figure 3. KWOC Index with Terms Repeated

COBALT-60 IN BITCHES ON LOW AND MODERATE FAT INTAKE.	+	1	PSEB-0111-0074
YIELDS.	EFFECTS OF LOW DOSES OF GAMMA RADIATION ON FLAN		IJAR-0013-0407
ESIUM-24 (P.P/GAMMA) AT LOW ENERGIES.	+	MECHANISM IN NAGN	PHRV-0120-2261
INELASTIC SCATTERING OF LOW ENERGY PROTONS FROM NEON-20.			MUPH-0039-0000
RUS-31 (P.GAMMA) SULFU	LOW ENERGY RESONANCES IN THE PHOSPHO		MUPH-0037-0031
GLANDS.	EXTRACTION OF LOW MOLECULAR WEIGHT CORTICO TROPHIC		ACEM-0042-0000
ID FROM C.	ISOLATION OF LOW MOLECULAR WEIGHT RIBO NUCLEIC AC		SCIN-0020-0400
RIZATION OF ETHYLENE AT	LOW PRESSURES.	METHOD FOR POLYME	ANCE-0074-0077
CORTX AS A CAUSE OF	LOW RESISTANCE OF THE NEWBORN AND		PEST-00-00-000
N THE	+	INFLUENCE OF A	LOW SODIUM DIET ON DYNAMIC CHANGES I
MELINES WITH CAUSTIC	+	LOW TEMPERATURE AGGLOMERATION OF NEP	IVUT-00-00-003
AND	+	OXIDES FROM THE	LOW TEMPERATURE OXIDATION OF NIOBIUM
FORMED ALUMI	NATURE OF	LOW TEMPERATURE TRANSFORMATIONS IN O	FMMY-0014-0750
OF LINEAR POLYMERS AT	LOW TEMPERATURES.	VIBRATION SPECTRU	DANK-0147-0500
ETIC BALANCE FOR USE AT	LOW TEMPERATURES.	SOUV NAGN	BOSJ-0030-2040
ON ADMIXTURE NUCLEI AT	LOW TEMPERATURES.	MOSSBAUER EFFECT	DANK-0147-0500
MUATION IN MAGNESIUM AT	LOW TEMPERATURES.	ULTRA SONIC ATTE	JUPS-0017-1001
DIOLYSIS OF N-HEXANE AT	LOW TEMPERATURES.	+	REACTIONS IN NA
NITRIDE INCLUSIONS IN	LOW-CARBON HIGH-CHROMIUM STEEL.		IANN-02-00-073
MANGANESE REFINED FROM	LOW-CONCENTRATION ORES CONTAINING		ARRM-0007-0301
IPITATES	ESTIMATION OF	LOW-DENSITY LIPO PROTEIN IMMUNO PREC	CLCH-0000-0010
OF NEGATIVE HYDROGEN	+	LOW-ENERGY COLLISION CROSS SECTIONS	JCPS-0037-2071
LEVELS OF MOLECULES BY	LOW-ENERGY ELECTRON IMPACT SPECTROSC		JCPS-0037-2497
ENERGY DISTRIBUTION IN	LOW-ENERGY ELECTRON-PHOTON SHOWERS		PHRV-0120-2352
ECTRUM OF INDIUM.	LOW-FREQUENCY LATTICE VIBRATIONAL SP		JCPS-0037-2727
WIGNIC CONVERTER WITH	+	LOW-FREQUENCY OSCILLATIONS IN A THER	CRAB-0010-0000
GANIC COMPOUNDS USING A	LOW-LEVEL NEUTRON SOURCE.	+	IN OR
DEUTEROLYSIS RATES OF	LOW-MOLECULAR-WEIGHT ALKANE SULFONYL		BOSJ-0030-2040
HYDRIDE.	POSSIBLE	LOW-PRESSURE POLYMERISM IN LITHIUM	ZPCL-0221-0100
OF TYLENOL OIL AFTER	LOW-TEMPERATURE CATALYTIC TREATMENT.		JCPS-0037-2730
WITH DI CHLORIDES FOR	LOW-TEMPERATURE DEWAXING OF OILS.		CRAB-0010-0000
MYLENE.	LOW-TEMPERATURE POLYMERIZATION OF ET		KTTN-07-12-011
			ANCE-0074-0000

Figure 5. Effects of Hyphen Usage Upon Computer Sort

One of each appears under "low pressure." Following this is one of the most common: "low temperature," with eight unhyphenated entries to three hyphenated--an overwhelming verdict for abandoning the hyphen in this case.

An additional factor for consideration is that most computer sorting treats a hyphenated word as one word--that is, where "low temperature" appears as two separate words, the citation will appear under both "low" and "temperature." Where the hyphen is used, the citation will appear under "low," but not under "temperature."

Such details as the use of singular and plural word forms also can cause difficulty (Figure 6). "Line" and "lines" have been separated by seven entries of "linear" and "linearity." This may seem insignificant in the example, as shown, and it is, because the eye can scan enough of the page at one glance

<p>FERTILIZATION. LINE AND P PLACEMENT EFFECTS ON REACTIONS IN HEATED LINE-ALUMINA MIXTURES.-</p> <p>EFFECTS OF FERTILIZERS. LINE AND CULTIVATIONS ON YIELD. ATION IN THE PROBLEMS, LIMITATIONS, AND THE FUTURE OF AUTON METRY.-</p> <p>DETECTION LIMITS IN RADIATION AND OPTICAL PYRO WITHIN THE TEMPERATURE LIMITS 20-60-DEG.-+ GLASSY TEXTOLITE</p> <p>LEAR MAGNETIC RESONANCE LINE OF ZEOLITIC WATER.-+ OF THE NUC</p> <p>MAGNETIC RESONANCE WIDE LINE SHAPES GENERATED BY TWO BROADEN</p> <p>MAGNETIC RESONANCE WIDE LINE SPECTRA.- ANALYSIS OF NUCLEAR</p> <p>N IODIDE.-</p> <p>MAGNETIC RESONANCE LINE WIDTH IN GARNET AND SPINEL TYPE</p> <p>ORGANIC ELECTROLYTES IN LINEAR AND CIRCULAR CHROMATOGRAPHY</p> <p>N-RANDOM DEGRADATION OF LINEAR CHAIN MOLECULES.- + OF A NO</p> <p>N-RANDOM DEGRADATION ON LINEAR CHAIN MOLECULES.- + ON THE NO</p> <p>THE ANTI FERRO MAGNETIC LINEAR CHAIN.- + SPECTRUM OF</p> <p>DER-DISORDER THEORY FOR LINEAR COLLOIDS.- ON</p> <p>OF THE COEFFICIENT OF LINEAR EXPANSION OF GLASSY PLASTICS.</p> <p>OF THE SPECIFIC HEAT OF LINEAR POLYMERS AT LOW TEMPERATURES.</p> <p>INVESTIGATION ON THE LINEARITY OF THE TEST CURVE FOR THE</p> <p>CHANGE OF K(3S)- LINES AND THE VALUE OF THE INITIAL</p> <p>RY OF SPECTRAL HYDROGEN LINES IN A PLASMA.- ASYMMET</p> <p>EUROPIUM(III) EMISSION LINES IN EUROPIUM DI BENZOYL METHANE</p> <p>SITES OF THE MANGANESE LINES.-+OF THE DIFFERENCE OF THE DEN</p> <p>MMA-CONJUGATE SYSTEM OF LINKAGE.-+COMPOUNDS WITH A CLOSED CA</p> <p>CYANATE ON ETHYLENIC LINKAGES. ANALYTICAL APPLICATIONS.</p> <p>INHIBITION OF ENERGY-LINKED DI PHOSPHO PYRIDINE NUCLEOTID</p> <p>ODD NUMBER OF MONOMERIC LINKS, FORMED IN THE THERMAL POLYMER</p> <p>ODUCT+ISOLATION OF PURE LINOLENATE AS ITS MERCURIC ACETATE A</p> <p>E+UREA FRACTIONATION OF LINSEED OIL FATTY ACIDS. COMPARATIVE</p> <p>CTIVITY OF LIPO PROTEIN LIPASE IN VARIOUS TISSUE SLICES.-+ A</p> <p>IGGER ACTION OF PHOSPHO LIPASE ON A MAST. CELLS.- TR</p> <p>PASE IN + LIPO PROTEIN LIPASE. ACTIVITY OF LIPO PROTEIN LI</p> <p>LYCERIDES BY PANCREATIC LIPASE.- + ENZYMIC HYDROLYSIS OF G</p> <p>F INTestinal + PHOSPHO LIPID COMPOSITION AND TURNOVER IN RA</p> <p>LIPID COMPOSITION OF TUMOR CELLS.-</p>	<p>SSSA-0020-0074</p> <p>JACH-0012-0030</p> <p>JSSC-0013-0021</p> <p>ANYA-0102-0171A</p> <p>JOSA-0052-1307</p> <p>PLNS-02-11-003</p> <p>CONE-0250-3400</p> <p>ACSA-0010-2212</p> <p>ACSA-0010-2149</p> <p>JCPS-0037-2699</p> <p>FVVT-0004-3604</p> <p>CANU-0007-0000</p> <p>JUPS-0017-1014</p> <p>JUPS-0017-1094</p> <p>PHRY-0120-2131</p> <p>JCPS-0037-2323</p> <p>PLNS-02-11-036</p> <p>DANK-0147-0500</p> <p>ZACP-0192-0370</p> <p>FMNT-0014-0400</p> <p>DANK-0147-0342</p> <p>JCPS-0037-2333</p> <p>MSKF-0000-0023</p> <p>DANK-0147-0630</p> <p>CHAL-0004-0403</p> <p>JACH-0230-0440</p> <p>DANK-0147-0106</p> <p>JAC-0030-0517</p> <p>IJAC-0025-0003</p> <p>IGSB-0004-0100</p> <p>AIPT-0140-0077</p> <p>IGSB-0004-0100</p> <p>MNKK-0030-1004</p> <p>NATU-0197-0070</p> <p>SIJO-0004-0200</p>
--	---

Figure 6. Effects of Singular-Plural Usage Upon Computer Sort

to see both entries even though they are separated. If this index covered 5,000 to 6,000 items, these terms could appear so far apart that the searcher would not see them both, and thus would miss a portion of the indexed information.

Many of the titles used today have meaningful terminology because the authors were aware that the titles were likely to appear in KWIC or KWOC index formats. But what becomes of a title such "A Look Ahead," or "Structures and Economics," or "Pilot Studies and System Components Evaluation"? Each of these papers is concerned with means of handling information, and each was presented at a national symposium on Engineering Information, the theme of which was "A Coordinated Engineering Information System." The difficulty

in indexing titles such as these shows that even information specialists are not making their own jobs easier. This is why librarians must enlist the help of societies and associations such as STWP in reminding others, and remembering themselves, that each piece of documentation is not a unique bit of printing destined to exist forever in its own little cubicle. The present store of documentation is so vast that it is almost incomprehensible, but every new publication must be added to it after the first use. To provide recall from this mass, the major subject or subjects of individual documents must be indicated to the indexer, whether this indexer is a machine or a man.

Examination of the growth of the application of the key-word indexes reveals their importance and indicates the stress that must be placed upon them. In 1958, H. P. Luhn proposed the KWIC concept. Later in that year, Chemical Abstracts Service received a National Science Foundation grant to apply Luhn's concept to chemical information. By 1962, over 30 different permuted indexes were in use, and by 1965 the number had grown to several thousand. The current indexes include such products as "Chemical Titles," published by the American Chemical Society, "Index to Legal Theses & Research Projects," published by the American Bar Foundation, "Current Medical Terminology," published by the American Medical Association, and "Meteorological and Geostrophysical Titles," published by the American Meteorological Society. An announcement received recently by the Douglas Library listed permuted indexes available to technical publications such as the Military Electronics Conferences, National Telemetering Conferences, Institute of Radio Engineers, and Institute of Electrical and Electronic Engineers' Transactions, Environmental Sciences Proceedings, and others. Many more of these indexes have been developed for use in industry. Many are produced for internal use only, as, for instance, at Douglas. The extent of the influence of these indexes must be recognized, whether or not the form is accepted as a proper index.

Although the indexing requirement of a title should not take precedence over its main purpose--to describe the contents of the document--nor be considered the only important feature to an indexing system, its value in automatic indexing cannot be overstressed. The Abstract, the Table of Contents,

the Introduction, and the Conclusion also are relied upon to provide guidance for the indexer under almost all circumstances. However, under conventional cataloging conditions, the human indexer is able to exercise some judgment with regard to the relative merits of terms, descriptors, subjects, and so forth. The major emphasis here is directed towards the mechanically created index, where human judgment cannot be applied once the indexing program is established. So even those who question the value of permuted indexes should provide titles that can be used.

Another serious problem arises in relation to the publishing of all these papers and documents. Every author enjoys having his creation appear in print, and without qualification, "the more the better." Certainly, each publication is directed toward a different audience and justification for duplication is easy. But what does duplication do to the cataloger, the reference librarian, and the researcher? Imagine the frustration of finding five citations to a subject, only to discover that the five citations cover two articles published in five different sources! Such multiple publication of the same article causes problems for the scientist and engineer as well. Also, the engineer's sources or references may be credited to a corporate author or publication different from that used by the library for the same article.

Figure 7 shows an example of the many corporate authors possible for a single article. For this particular item, the work was done by a member of the staff of the California Institute of Technology, under funding from the Air Force Aerospace Research Labs, and submitted for publication to the "American Journal of Physics." The report was transmitted per contractual agreement to the Defense Documentation Center for dissemination throughout the defense community. Because of its application to aerospace endeavors, the National Aeronautics and Space Administration also acquired the document for dissemination. Thus, there are five sources established, three of which are primary and two (DEC and NASA) are secondary.

Are not all of these sources entitled to claim this document as their own? Certainly they are. The Aerospace Research Lab paid for the research, the California Institute of Technology prints its staff's output, the "American

PDL 56129

CALIFORNIA INST OF TECHNOLOGY TR 32-738
JET PROPULSION LAB
AEROSPACE RESEARCH LABS ARL 65-151
AMERICAN JOURNAL OF PHYSICS
/DEFENSE DOCUMENTATION CTR/ AD 622 277
/NATIONAL AERO AND SPACE ADM/ N65-31848

PLASMA NONUNIFORMITY AND GRID PLASMA
EROSION IN AN ELECTRON BOMBARD- NONUNIFORM
MENT ION ENGINE. 7P. GRID

D. J. KERRISK, T. D. MASEK

30 JULY 1965

C.1, C.2

EROSION
ELECTRON
ION
ENGINE
ELECTRON-
-BOMBARDMENT
EMISSION
SURFACE
DENSITY
FLOW
RATE

Figure 7. Index Card with Multiple Corporate Authors

Journal of Physics" agreed to publish the article for circulation among the scientific community. Both DDC and NASA are obligated under their charters to announce and disseminate all work performed through government funding.

Those responsible for the publication of material cannot solve this problem by limiting the number of times a report is published, but they can ensure that each publication makes note of the other sources in which it appears. When previous or prospective publications are known, why should not this information be included in a footnote? If the specific report number or issue number were not available, it would be helpful if a statement such as this were used: "Also published as Jet Propulsion Lab Technical Report and Aerospace Research Lab Technical Report," or "Announcement is anticipated in the Defense Documentation Center's Technical Abstract Bulletin and the National Aeronautics and Space Administration's Scientific and Technical Aerospace Reports Bulletin." Either of these statements would provide both cataloger and scientist with a warning that the document may appear under any one of those various corporate authors.

Often library users do not recognize the same material in different forms. Having once seen a journal article, a user may not be aware that the grey-cover ARL Report contains the same material--at least, not until he has requested the document and has begun re-reading it. This can also be true of the cataloger and librarian. Yet, with the tremendous quantity of information published today, information specialists must do all they can to save the reader from such circumstances. Libraries are given generous budgets, professional staffs, and expensive computer programs to inform their clients of material of interest to them. Informing an individual over and over again of the same material is not pleasing to either the individual or the responsible information specialist. The integrity of the entire information handling system is downgraded by these occurrences. To help avoid this, the author and the technical writer should give notice of the multiple publication of any documents he handles.

With the widespread use of mechanized systems has come another problem--the variations used in individual corporate names. Some mechanized programs have avoided the problem to some degree by coding the corporate authors. However, this imposes another set of problems. The concern of the information specialist is with the use of company names, agency names, universities, and so forth, as they are input to a computer. Take, for example, "IBM." In most communications, the use of "IBM" conveys the desired message. But the name of the company is International Business Machines, so corporate source notations should appear with the full name, "International Business Machines." This is a simplified example, because most catalogers would know that the citation should use the full name. It is not always so obvious, however, whether initials can constitute an acceptable corporate entry. IBM is not acceptable, because this is not the legal name of the organization. IIT Research Institute is acceptable and correct, since this organization officially discontinued the use of the title "Illinois Institute of Technology Research Center" and adopted IIT Research Institute as its official name. These examples illustrate how confusing to new and

inexperienced catalogers abbreviations in corporate names can be. Such abbreviations should be avoided unless they constitute a legal designation.

Even such variations as abbreviating "company" one time and spelling it out the next time can cause confusion. When a computer sorts, it sorts according to standard computer rules. Douglas could appear at least four different ways (Figure 8), without counting eliminating the periods or including divisional entities. The computer does not recognize all these names for the same company--it recognizes only a change in the characters in each.

Authors, technical writers, and information specialists must all be certain to use the corporate name in the same manner each time, so that the items will all appear together in a computer printout. It is important for indexers and catalogers to be precise about entries, but the tendency is to accept an

DOUGLAS AIRCRAFT COMPANY
DOUGLAS AIRCRAFT CO.
DOUGLAS AIRCRAFT COMPANY, INC.
DOUGLAS AIRCRAFT CO., INC.

Figure 8. Variations of Single Corporate Author

entry as it appears on the document. For this reason, the careful use of the corporate name can be of great help in facilitating the retrieval of valuable information.

Projections into the future indicate that, without doubt, the next major breakthrough in information systems will be total text input of some sort. Paper tapes and optical scanning methods have been tried with varying degrees of success but they reveal that there are several obvious questions to be answered in relation to full text input. These are as follows:

1. What about the extent to which the "stop list" would have to be enlarged beyond that of the KWIC stop list?
2. What about synonyms and near-synonyms? In the interest of variety, the writer will strive for a variation in terms, perhaps using words such as automobile, car, vehicle, sedan, all in reference to the same subject. When retrieving from total text input, care must be taken to ensure that each of these terms is used for recall.
3. What about antonyms? When a request arises for information on safety in handling chemicals, will there be a reference also to "hazards in handling chemicals"?
4. What about adjectives and adverbs that lose meaning when taken out of context? These words create problems, not so much in retrieval as in clutter (Figure 9). They are a vital part of all writing for logical communications, yet they are rarely used for retrieval. One interesting exception is the term "orbital." When first entered into the Douglas dictionary, the only term used was "orbital," although the usual procedure was to enter both the plural and singular forms of all terms. It seemed obvious that no one would ever use "orbital" by itself, much less in plural form. Yet within no more than 6 months, Douglas librarians were not only coping with "orbital paths," as expected, but also with "orbitals." The plural form has now become an accepted term in the aerospace industry.

KEY WORD LIST

1. IN	DERIVE	HIGH	PROBABLY
1	DERIVED	IN	PROBLEM
29	DESCRIBE	MALL	PROCESSING
2	DESCRIBING	IND	PROCESS
3	DETERMINE	INDICATING	PROFILE
4	DETERMINING	INDEXED	PROJECT
5	DIFFICULTY	INDEXER	P
ACCORDING	DISTINCT	INDEXY	RAT
ADDITIONAL	DISTINCTIVE	INDEXING	REFERENCE
ADDITION	DOCUMENTALIST	INDEXES	REFERENCE
AMERICAN	DOCUMENTATION	INDICATED	REFERENCE
AMONG	DOCUMENT	INDICATE	REFERENCE
APPEARED	ELIMINATED	INDICATOR	REFERENCE
APPEARING	ELIMINATE	INFORMATION	RELATION
APPEAR	ELIMINATING	INTERESTED	RELATIONSHIP
APPLICATION	ENGINEER	IRVING	RELEVANCE
ARRANGED	EXAMPLE	LANGUAGE	RELEVANT
ARRANGE	EXISTING	LIST	REPORTED
ASSOCIATED	EXPANDED	LITERALLY	REPORT
ASSOCIATION	EXPERIENCE	LITERATURE	REQUESTED
AUTOCODED	EXPERIMENTATION	LOGIC	REQUESTOR
AUTOENCODING	EXPERIMENT	LUMP	REQUESTING
AUTOMATICALLY	EXPOSURE	MACHINERY	REQUEST
AUTOMATIC	FACT	MACHINE	REQUIRED
BASED	FACTOR	MARCH	REQUIREMENT
BELIEVED	FILM	MATERIAL	REQUIRED
BELIEVE	FIRST	MEANINGFUL	RESEARCH
B	FOLLOWING	MEANING	RESULTANT
CALLED	FOLLOW	METAL	RESULT
CLOSELY	FORM	METHOD	RETRIVAL
COLLECTION	FORMULA	NEAR	SEARCHED
COLUMN	FOUND	NEW	SEARCHING
COMPARED	FREQUENCY	NUMBER	SEARCH
COMPARE	FREQUENTLY	OCCURRED	SECOND
COMPARING	FREQUENT	OCCURRENCE	SELECTED
COMPUTATION	FRICTION	ONE	SELECTING
COMPUTER	F	ORDER	SELECT
COMPUTING	FURTHER	ORIGINALLY	SEMANTICALLY
CONSIDERED	FUTURE	ORIGINAL	SEMANTIC
CONSIDERING	GENERAL	PREPARATION	SINGLE
CONTAINED	GENERATED	PREPARED	SMALL
CONTAINING	GENERATE	PRESENTATION	SPECIFIC
CONTAIN	GENERATING	PRESENT	SPECIFIED
COORDINATE	GENERATION	PRESERVATIVE	SPECIFY
CORDELL	GIVEN	PROBABILISTIC	STAT
C	GROUP	PROBABLE	STATISTICALLY
DEGREE			
DERIVED			

Figure 9. Automatic Index

CONCLUSIONS

The technical writer's key to computerized information retrieval is not the restriction of all writing to a 10,000-word vocabulary. Rather, the key is the thoughtful selection and use of truly pertinent and descriptive terms, especially in those portions of a report which serve as key tools for effective storage and retrieval of valuable information, and in fully documented publication notations. Those who feverishly embrace the new word, or the new concept described by an old word, must consider the problems of the information retrieval specialists, and must realize that the kind of vocabulary used today must be retrievable tomorrow.

BIBLIOGRAPHY

M. Fischer. The KWIC Index Concept: A Retrospective View. American Documentation, April 1966, Pages 57-70.

A. G. Hoshovsky. How Technical Writers Can Use and Improve Technical Retrieval Systems. Office of Aerospace Research, 17 June 1966.

M. E. Stevens. Automatic Indexing: A State-of-the-Art Report. National Bureau of Standards, NBS Monograph 91, 30 March 1965.

Chemical Titles. American Chemical Society, No. 26, 19 December 1966.

Second National Symposium on Engineering Information. Engineers Joint Council, March 1966.

Symposium on Materials Information Retrieval. Aeronautical Systems Division, ASD-TDR-63-445, May 1963.