EFFECTIVENESS OF KWIC INDEX AS AN INFORMATION RETRIEVAL TECHNIQUE FOR SOCIAL SCIENCES*

By YOSHIRO MATSUDA**
AND
SACHIKO MATSUI***

Introduction

The literature search is one of the major aspects of the research activity and deserves the researchers’ serious consideration and yet has attracted least attention except that of the documentalists and librarians. But recent increase of the publications and growth of the various means of information have brought up the mechanized information processing in literature search first in the field of natural sciences and gradually in the field of social sciences.¹

Recent development of the computer has made a revolutionary change to the processing of the non-numerical information. The powerfulness is overlooked or not fully realized among the researchers in the field of social sciences; it is so even among the econometricians whose daily work heavily depends on computer and even more among the historians of economic and social thoughts whose hostility against computerized information retrieval is quite strong.

Our survey, though limited in scope, shows that several important trials of computerized information retrieval in action in the field of social sciences have been ventured mostly by the social scientists of particular disciplines without or with least collaboration of librarians.² Thus, the mechanized literature search is faced with the indifference both from users and librarians. This paper aims to fill the credibility gap between the researchers of documentation and researchers of social sciences through presenting the main findings

---

* This work was carried out by the group conducted by Yoshiro Matsuda with the aids of the grants from Isetan Scholarship Foundation for 1970-71 and the Ministry of Education for 1972-73 (Experiments for comparison of effectiveness of retrieval by KWIC index and facet classification scheme using the titles and abstracts.) The authors wish to thank to all the participants to the project for their stimulating discussions and co-operation and for their generosity to use whole results and findings freely in this paper; Professor Jun’ichi Akiba of Hitotsubashi University, Assistant Professors Shigeyuki Sato of Hokkaido University and Ryo Suzuki of Saga University and Lecturers Akira Hirota of Hosei University and Kiyonari Kidahashi of Hokkai Gakuen University. Computer processing is executed at Computing Centre of Hokkaido University.

** Assistant Professor (Jokyōju), Documentation Centre for Japanese Economic Statistics, Institute of Economic Research, Hitotsubashi University.

*** Assistant, Department of Management Sciences, Faculty of Commerce, Otaru University of Commerce.

¹ Janda [10] especially Part II; Research applications, Matsuda [14].

² More exactly without the initiative of traditional librarians or bibliographers. See the Table in Appendix 1.
of our trials published in the *KWIC index series for social sciences*,3 comparing them with other similar trials in the field of social sciences.

Such advanced documentation activities undertaken in a specified discipline as abstract and / or index making have often encountered with suspicion against their usefulness. Such suspicion shares the following two suppositions.

1) In the case of social sciences, literature search is so unique and peculiar to the researcher's personal frame of reference or value premises that any information processing planned for common users is quite unpractical for the researcher. And librarians are expected to remain as a neutral interpreter to writers and users. This "librarians' neutrality" is guaranteed so far as he depends on usual devices of information control such as classified catalog or subject catalogs in a traditional form.

2) It is for the current awareness services and not for the retrospective search which is indispensable for social sciences depending on the search of so huge pile of volumes stored from the past that the recently developed information retrieval techniques are most suitable.

The first proposition is less supportable judging from the various surveys which show that the librarians' classified and subject catalog themselves have come to be a jargon not only for the outsider researchers in other disciplines but also for the insiders like library science specialists or information officers.4

Severe attack comes from the researchers in the interdisciplinary sciences. For "the librarians' neutrality" is supported by the bias that the librarian's or indexer's judgement is itself neutral. But such neutrality can scarcely be supported; because the descriptors used by librarians, which are so to speak subject headings, require the writer to transform his intention into librarians' vocabulary and the user of these descriptors is in turn requested to transform his retrieving vocabulary into librarians' vocabulary. These double transformations will not guarantee the equivalence of the semantics of the writer's descriptors and user's descriptors and make it difficult for librarians to remain as a neutral interpreter or translator.

It is clear that in the case of classified catalog the same problem occurs more severely. It should be well remembered that Bliss classification group and CRG inspired by Ranganathan's theory of faceted classification have gained practical success in particular subject area.5 This suggests that universal classification scheme is difficult to be designed and single purpose classification scheme is much more powerful and, thus it may be implicitly supposed that librarians' non-neutrality or the homogeneity among the writers, librarians or indexers and users. From this it will be well supposed that librarians' efforts could be flourished when they commit in a particular discipline or a frame of reference. It is this kind of reasoning that stimulates the direct descriptor making tendencies represented by the KWIC or KWOC indexes now in fashion.6

---

3 See [B-6].
5 T.D. Wilson; "The work of the British Classification Research Group," in Wellisch [24]. They still hold the possibility of compiling the effective universal classification scheme.
6 As to the development of these techniques, see Fischer [7]. Another approach of compiling descriptors eliminating the indexer's bias, see Jean Aitchison; Thesaurofacet, a new concept in subject retrieval schemes," in Wellisch [24].
However, the mechanical information processing is undertaken in most cases through the collaboration of the researchers of a particular discipline and documentalists or computer specialists, excluding the librarian's participation. These facts show that a philosophical change within library science is required even when it is applied to the field of social sciences. Thus we should examine the plausibility of the second proposition.

Table 1 shows the chronological development of KWIC index applied to information retrieval in social sciences and natural sciences. The time lag of the impact of computerization is not large between them but the stage of practical application comes so soon in the field of natural sciences while that in social sciences is staggering in the experimental stage.  

But, even from this short table, it is clear that the KWIC index compilation is more favoured in the interdisciplinary sciences like sociology, cultural anthropology, political science, econometrics and computer sciences or in the field where like legal processes immediate informations are required. Thus this kind of information retrieval techniques has a bright future in the interdisciplinary sciences. The most important point is that trials in the social sciences are mostly oriented to the retrospective research and not for current awareness service. The prosperity of the cumulative indexes may be its evidence and acts also as a counter example for the first proposition.

Needless to say there remains a supporting evidence for the second proposition. The past publication of mankind is so huge that even wealthy United States cannot afford enough financial support for her RECON project; to convert her national union catalogue into machine readable form. But there are two examples qualified enough to support the retrospective search in social sciences. One is Harvard University's Widener Library Shelflist Conversion and Publication Program resulted in the publication of the Harvard Widener Library Shelflists which are facsimile of computer outputs having author, title and chronological lists in addition to the main list (classified shelflist). This series promises us the future fertility of the retrospective research brought out through the further corporation of the important libraries sharing the MARC (machine readable cataloging) program.

The second example is Pollin's Godwin Criticism which covers about 3,379 items published in 1783-1966. Although it has some defects which we will discuss later, it has broken a new path for the retrospective search by machine. And we might be permitted to register our work as the third example in this field. It deals with about 686 works of Claude Henri de Saint-Simon and Saint-Simonians published in 1817-1971.

I. Natural Language or Descriptor

Here we will analyze the possibility of neutral transformation of the contents by virtue of the librarians' descriptors called Subject Headings (hereafter often abbreviated as SH). The cease of publication of Jurindex occurred in the second year of its short life is quite symbolic and it symbolizes the present situation of the mechanical processing in the field of social sciences. A few periodical publication of the mechanized indexes are established like the Annual supplements to the Universal Reference System; Political Science, Government & Public Policy: an annotated and intensively indexed compilation of significant books, pamphlets, and selected and processed by the Universal Reference System—a computerized information retrieval service in the social and behavioral science. 1965+

De Gannaro [5].

Pollin [B-5].
<table>
<thead>
<tr>
<th>Year</th>
<th>A. INDEX &amp; ABSTRACTS</th>
<th>B. CUMULATIVE INDEX</th>
<th>C. SUBJECT BIBLIOGRAPHY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1962</td>
<td>Current Legal Legislation Index, (American Bar Foundation)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1967</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1970</td>
<td>JURINDEX, l'informatique appliquée à l'actualité bibliographique du droit francophone. 1971–1972. KWIC (Institut de recherche d'informatique juridique.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1971</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1973</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Bwy, Douglas; Social Conflict; a keyword-in-context-bibliography on the literature of developing areas, with supplementary reference from Latin America.

Aldous, Joan, & Hill, Reuben; International Bibliography of Research in Marriage and the Family. 1900–1064.

Matsuda, Yoshiro & Matsui, Sachiko; Bibliography; Works on econometrics, reviewed in Econometrica. 1951–1971.

Matsuda, Yoshiro, Sato, Shigeyuki Hirot, Akira & Matsui, Sachiko; Bibliography; Works by Saint-Simon and Saint-Simonians. 1802–1970.
as SH). They are expected to be called neutral to both writers and users but in fact, should be regarded as to be subject to the librarians' frame of reference or paradigm. The most important paradigm of librarian to compile SH is "public use first, specialist use second". Thus users sometimes encounter with a serious gap between technical terms which now become of their common usage and SH terms using colloquial and sometimes unsophisticated words which are seldom used in researchers' discussion. This fact must have been stemmed from the belief that the paraphrasing by daily language will help the recognition of the public easier. This belief is severely criticized by John M. Christ who takes the SH of the Library of Congress of the United States as an example of the most popular SH.\(^{10}\)

He compared the subject headings of social sciences in the 7th edition and the Supplements of the Library of Congress Subject Headings (hereafter abbreviated as LCSH) with

### Table 2. The Discrepancy of the Usages of Technical Terms among Library Scientists and Social Scientists\(^*\)

<table>
<thead>
<tr>
<th>Terms that match with LCSH</th>
<th>Main index headings of IESS(^{**})</th>
<th>Interdisciplinary terms(^{***})</th>
<th>Sociological terms(^{****})</th>
</tr>
</thead>
<tbody>
<tr>
<td>280 (42.5%)</td>
<td>24 (19.50%)</td>
<td>163 (34.5%)</td>
<td></td>
</tr>
<tr>
<td>145 (21.5)</td>
<td>24 (18.50%)</td>
<td>53 (11.5)</td>
<td></td>
</tr>
<tr>
<td>84 (12.5)</td>
<td>19 (14.50)</td>
<td>33 (7.5)</td>
<td></td>
</tr>
<tr>
<td>156 (23.5)</td>
<td>61 (47.5)</td>
<td>225 (46.5)</td>
<td></td>
</tr>
<tr>
<td>Total terms examined</td>
<td>665 (100%)</td>
<td>128 (100%)</td>
<td>470 (100%)</td>
</tr>
</tbody>
</table>

\(^*\) Figures are based on J.M. Christ's survey. See Christ [3], pp. 68, 83 & 123. A miscalculation of percentage is corrected. Comparison is based on the Subject Headings of Library of Congress (LCHS). 7th edition. \[22\].


\(^{***}\) Interdisciplinary terms are based on the terms used in the 13 basic interdisciplinary social science texts. See Christ [3], pp. 74–5. Whole list of these 128 terms are listed in table 5 of Christ [3], pp. 76–81.

\(^{****}\) Sociological terms here defined are extracted from the indexes of Broom Leonard & Philip Selznick's *Sociology; a test with adapted readings*, Evanston, Ill., 1955 (287 terms) and Robert Faris (ed.)*'s *Handbook of Modern Sociology*, Chicago, 1964 (235 terms), sharing only 41 terms in common.

\(^{10}\) Christ [3]. As to the criticism and defense to the LCSH, see Richard S. Angell; "Library of Congress Subject Headings—Review and Forecast," in Wellisch [24].
the index headings from the *International Encyclopedia of the Social Sciences*. His comparison was carried out as to i) general social science terms, ii) interdisciplinary social science terms and iii) sociology terms and other detailed comparison of particular concept like “value”. His conclusion summarized in Table 2 was that LCSH has governed by different frame of reference from that of the social scientists. A quarter of general terms cannot be traced and in the case of sub-terms having the form of compound terms untraceable ones amount to 93.3%. The LCSH prefers single word heading contrary to the tendency among the scientists to prefer corporate words or collective terms. This is shown in Table 3.

Table 3. Characteristics of the Subject Headings Shown in the Number of Words in the Subject Headings

<table>
<thead>
<tr>
<th>Terms without subject heading referents</th>
<th>Terms with subject heading referents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single word</td>
<td>Multi or corporate words</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1) General social science terms in the main headings of IESS (partial)</td>
<td></td>
</tr>
<tr>
<td>14 (27%)</td>
<td>38 (73%)</td>
</tr>
<tr>
<td>2) Interdisciplinary social sciences terms (partial)</td>
<td></td>
</tr>
<tr>
<td>22 (49%)</td>
<td>24 (51%)</td>
</tr>
<tr>
<td>3) Sociological terms (partial)</td>
<td></td>
</tr>
<tr>
<td>12 (24%)</td>
<td>39 (76%)</td>
</tr>
<tr>
<td>4) Management sciences terms (LCSH before &amp; in 7th ed.)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>5) Management sciences terms (after 1966 or Suppl. to 7th ed.)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>6) Sub total of 4) &amp; 5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sources: From 1) to 3) figures are obtained from Christ [3] pp. 68, 69—70, 83—85, 121—3. But some counting errors in the original tables are corrected. From 4) to 5) figures are obtained from Matsui [15].

To interpret his finding we must take into consideration the following: that articles or index terms in Encyclopedia treat narrow concepts like articles in academic journals or in one chapter of a book but that books on which LCSH is assigned are apt to treat much broader concepts. But his findings are still striking enough. And this must be the background of many trials of thesaurus construction.

---

For LCSH is also slowly catching up this process. In the 5th edition of LCSH one thirds of the LCSH are single word and the 7th edition shows increase of corporate or multi words. See Daily [4] p. 3962. And the Table 3 shows that this tendency is still going on after the 7th edition.
**Fig. 1. Examples of KWIC Index**

1. From [B-6, no. 1] (LP output facsimile)

<table>
<thead>
<tr>
<th>KWIC column</th>
<th>Item number</th>
</tr>
</thead>
</table>

2. From [B-6, no. 2]


**1973 Effectiveness of KWIC Index as an Information Retrieval Technique for Social Sciences**

21
3. From [B-6, no. 3]

S. SIMON ET DÉ, /SAINT-SIMON UND DER
NO NEUFREI KOMPANISUS, VOL. 2. /SAINT-SIMON UND DER
MOISELLE /SAINT-SIMON UND DIE /SAINT-SIMON UND DIE
N HA/Saint-simon und jünges Deutschland, 1949. /SAINT-SIMON UND DIE
ALTHICHE, EINE SYSTEMATISCHE /SAINT-SIMON UND DIE
DESSON, SOCIALES WISSENSCHAFTSCHRIFTEN. /SAINT-SIMON UND DIE
/J/ /LA GNEVE DE
/JO //LA DOCTRINA
ALLA UCCHWA CATTOLICA, SIMPLEMENTE AL RICCGOL/DI
/CESARE //F //LA DOCTRINA
SCHE DARSTELLUNG DER IDEEN /SAINT-SIMON UND SEIN
ECÖDE D'UN CUPP, INTRODUCTION A L'ÉTUDE DE LA
/COMEE D'UN CUPP D'OEIL GÉNÉRAL, VIE L'ÉTAT DE LA
EUSE ENGAGÉE ENTRE LUI ET /SEILLES MINACLES ET LA
/LE COUNTER-REVOLUTION OF
/SUR LE MIHRYLITE CANNOT, MEMOIRE DE L'ACADÉMIE DES
MIHRYLITE CANNOT; MEMOIRE DE L'ACADÉMIE DES
CHEZ LES SAINT-SIMONIENS, THÈSE POUR LE DOCTEURATE
INT-SIMEN ET LA SOCIALESM, THÈSE POUR LE DOCTEURATE
LE SAINT-SIMONISME, COURS PROFESSÉ A L'ÉCOLE DES
/LES SOURCES FRANÇAISES DU SOCIALISME

4. From [B-6, no. 4] (KODAK, KOM-90 output facsimile)
His findings suggest us to use another kind of descriptors for subject retrieval. One extreme way is to use natural language used in the text to be retrieved. “Natural language” used here means vocabulary uncontrolled by indexers or librarians.

The most practical system of information retrieval using natural language may be the KWIC index system formulated by H.P. Luhn in 1959. As is well known KWIC means key-word-in-context index which uses title words as keywords instead of the thesaurus or SH except some stopwords preassigned and they are listed in alphabetical order in the middle of the page, accompanying the rest of the title after and before the keyword. See illustration in Fig. 1.

To compare the KWIC index with LCSH we should use the data base consisting of the bibliographic description of the documents in book form. According to our KWIC index to management sciences summarized in Table 4, the subject headings (LCSH) whose

| Table 4. Degree of Correspondence of Title Words to the Subject Headings of Library of Congress |
|-------------------------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Number of titles to which LCSH are assigned | Number of total LCSH assigned | Exact correspondence to title words | Partial correspondence to title words | No correspondence to title words |
| LCSh before & in 7th ed. | 907 | 1642 (100%) | 441 (27.1%) | 677 (41.2%) | 521 (31.7%) |
| LCSh after 1966 or suppl. to 7th ed. | 225 | 147 (100%) | 41 (27.9%) | 61 (41.5%) | 45 (30.6%) |
| Total | 1132 | 1789 (100%) | 485 (27.1%) | 738 (41.3%) | 566 (31.6%) |

Total different LCSH examined are 558 in 7th ed. and 77 after 7th ed. Some adjustments are carried over in matching the LCSH and title words and they amounts to three percentage of the total subject headings examined.


words and / or phrases share the title words in common are only 68.4%. And D.H. Kraft’s similar survey in the field of legal periodicals shows it 64.4%.

Although V. Mostecky pointed out that the subject headings should not share the words in common with the title words in order to avoid the vagueness found in the titles. But Christ’s figures shown in Table 2 indicate that through the subject headings almost 54% of the technical terms get no direct access and they must have escaped from Mostecky’s blame because they are more exact than the title words and belong to the researchers’ common language. Though Christ’s figures are not weighted with the frequency of occurrence in the published documents, they are quite close to those of ours and Kraft’s. This similarity suggests that the natural language or title words are much closer to the technical terms than LCSH in the

---

13 His pioneering article is now easy to access through the facsimile reprint in Schultz [21].
social sciences and that the utilization of the title words instead of the subject headings like KWIC index makes it easier to get access to the technical terms.  

Table 5 shows that the descriptors assigned to one title is about 6.2 on average in case of KWIC index and about 1.6 in case of LCSH. Thus the title words will bring out more access points than the usual subject heading approach.

As KWIC index requires no librarian’s work to examine the content of the book, the above comparison of natural language and LCSH might lead to a conclusion that the KWIC or mechanized information processing is superior to the subject heading approach.

But we should take the following points into consideration; the descriptiveness of the contents by the title decreases with the sophistication due to the writer’s literary taste. In this case some additional descriptors are required. This compensates the labour saving character of KWIC index making. Kraft’s experiment mentioned above says that 10.5% titles need this kind of correction and in K. Janda’s case on American Political Science Review it is 33%. Other experiment like Aldous & Hill’s case also reports necessity of the correction of this kind, though no statistics is reported.

Supplementary descriptors for KWIC index are required especially for the empirical research dealing with specific time and space elements. K. Janda’s work on political science, Aldous & Hill’s work on family and marriage and our work on econometrics will show a good example on this point.

In most cases, except ours, numerical figures are assigned as stopwords. The date in the title are excluded from the KWIC index and only verbal expression of time or period remains. However, the pronunciations of numbers and their spellings are so different even among European languages that diversification of the same word is too wide to be used for retrieval.

---

15 As to the application to the natural sciences, see Ruhe [20] and Rosenberg [19].
17 Except this point merger of different languages causes no serious difficulty in the case of European languages including Slavic. In case of Japanese the situation is rather different. Transcription via Roman alphabet is not applicable for Japanese, Chinese and Korean to be treated simultaneously.
Examining these previous results, we excluded numerical figures from the stopwords. In the case of econometrics, the dates shown in the title are concentrated to those in the twentieth century (about 40 titles) and the rest are 13 titles on the nineteenth century (including two covering the eighteenth century) and two exceptional titles: one is Gossen's biography 1810-1858 and the other is British historical analysis of 1790-1859. To obtain the information of these 55 titles occupying 105 lines, we should scan 308 lines in the KWIC index. Similar results are obtained in the cases of management sciences and others shown in Table 6.

<table>
<thead>
<tr>
<th>Total lines of numerical figures in KWIC index</th>
<th>Management Sciences</th>
<th>Econometrics</th>
<th>On Saint-Simon &amp; Saint-Simonians</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figures for the years and century</td>
<td>303</td>
<td>307</td>
<td>98</td>
</tr>
<tr>
<td>i) Description of the contents</td>
<td>90</td>
<td>115</td>
<td>70</td>
</tr>
<tr>
<td>ii) Century</td>
<td>78</td>
<td>6</td>
<td>70</td>
</tr>
<tr>
<td>iii) Date of the Conference, etc.</td>
<td>1</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Others (editions, etc.)</td>
<td>12</td>
<td>109</td>
<td>28</td>
</tr>
<tr>
<td>Total titles retrieved by i) &amp; ii)</td>
<td>212</td>
<td>192</td>
<td>40</td>
</tr>
</tbody>
</table>

In the case of LCSH, no exact dates are given even for the documents on history. The analysis especially on the contemporary phenomena requires exact specification of date and the figures in the title or in the sub-title are quite effective for this specification. But for this purpose, the assignment of such a date in the title is not sufficient and we should add some supplementary descriptors in figures.

As to assigning the area specification, LCSH is superior to the direct use of the title words, because in most cases those works written for writer's own countrymen seldom use his country's name in the title. See Table 7.

Thus additional descriptors for area designation are required for KWIC index used by the users aiming international comparison. Maybe for those whose primary concern rests on his native country such additions are unnecessary. For example K. Janda assigned AMERICA, AMERICANS, US as stopwords. His presupposition might be that articles in American Political Science Review deal mainly with American affairs. But when a bibliography becomes more international and extensive through inclusion of the articles or books written by foreigners, this kind of omission of the keywords becomes obstacle.

---

18 Our latest trial, library and information science seen in [B-6, no. 4], is requested to make a further improvement for its data base structure includes some figures useless for retrieving due to the inclusion of the contents of the documents like Encyclopedia of Library and Information Science. Thus the figures after V. for volume, P. for pagination, VOL. ED., NO., etc. are assigned as stopwords to be excluded from KWIC column. And figures like nineteenth century are transcribed into 19TH CENTURY. This technique is a variant of KWIC index. Similar trials are found in Matthews [17].

19 He did not make any explicit reasoning on this point. But he did not exclude UNITED STATES, U.S. and U.S.A. from keywords. We suppose that he needs UNITED, and U. to get UNITED NATIONS and U.N. as keywords. As to the case of Aldous and Hill no information on stopwords is given as mentioned before, this prevent us to evaluate this bibliography exactly. It should be remembered for the compiler of KWIC indexes that the exposition of the list of stopwords is necessary not only to show the compilation process but also to make a thorough retrieval.
for effective retrieval.

Our tentative conclusion is that KWIC index supplemented with dates and places will be a powerful substitute for the hand-made index like LCSH. And it may act as a more powerful retrieval technique with these supplementation and it will rarely meet incongruence between the librarians’ conceptualization and that by the social scientists.

<table>
<thead>
<tr>
<th>Place</th>
<th>Found</th>
<th>Not found</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Different</td>
<td>Total</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>19</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>33</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time</th>
<th>Found</th>
<th>Not found</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>17</td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Person</th>
<th>Found</th>
<th>Not found</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total</th>
<th>Found</th>
<th>Not found</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>66</td>
<td>55</td>
<td>66</td>
</tr>
</tbody>
</table>


The second point that should be taken into consideration is that, when the vocabulary is uncontrolled, the natural language of the same meaning will be scattered throughout the whole KWIC column. For example, VALUE THEORY—PRICE THEORY, INTERINDUSTRY ANALYSIS—INPUT-OUTPUT ANALYSIS—LEONTIEF ANALYSIS, etc. This problem will become more serious in the case of merging the different languages like WIRTSCHAFT—OEKONOMIE—ECONOMIE, or HISTORY—HISTOIRE—ISTORIA—GESCHICHTE, etc.

But as far as European languages are concerned our experiment shows remarkable success because their differences in spelling concentrate to the ending part showing inflexion, etc. This, however, cannot be applicable to Romanized items in Chinese, Korean and Japanese which share common letters in part.

Needless to say, it is impossible to retrieve all the relevant documents by KWIC index without thesaurus or controlled vocabulary lists. This defects cannot be driven away by means of SH approach because of the lack in hierarchical categories in SH. Thus we should make reservation for the final evaluation of classification scheme approach.50

---

50 See “Panel discussion,” in Wellisch [24] and the future possibility of information science suggested in the articles of Elias [6].
II. Retrospective Search and Content Analysis through KWIC Index

As mentioned above, interdisciplinary approach may become one of the most important aspects of our day and stimulate our interests in computerised indexing. And similar approach has occurred even in the field of analysis of social and economic thoughts. It is required to give the whole picture of one person who participated in several movements in the history of thoughts or one person whose activities and thoughts created a sort of a sect of social movements with international influence. Such a "whole picture" itself changes in accordance with the change of "paradigms." Thus the change of analysis itself becomes an object of research. These kinds of research work requires much more intensive bibliography than ever compiled. Recent close international contact among librarians and researchers has made it possible to compile such a bibliography in the worldwide scope.

To take for example, Anthony P. Campanella's *Giuseppe Garibaldi e la Tradizione Garibaldina: una bibliografia dal 1807 al 1970* includes about 16,141 documents written in 54 languages and dialects and Burton R. Pollin's *Godwin Criticism; a synoptic bibliography* covering 1783-1966 includes about 3,374 documents in 14 languages. The latter is compiled in machine readable form and provided with 12 computer made indexes and statistics tables. As far as we know, this may be the first trial of computerised bibliography designed for retrospective search for social sciences. But unfortunately this work sticks too much to the traditional form of a bibliography and does not realize the potential power of the machine readable form. Even in the field of humanities the computer processing of non-numerical information has come to be a strong weapon to compile a concordance or to identify the authorship of the anonymous or pseudonymous texts through quantification of the contents of the texts. Such a technique of analysis of non-numerical information may be applicable for the bibliography in machine readable form like *Godwin Criticism* equipped with short abstracts and many descriptors evaluating the degree of relevance to Godwin and/or his works. Abbreviations used for descriptor are A; article on Godwin, R; review of a book by Godwin, P; passage on Godwin, C; short comments, interspersed, M; mention of Godwin; B; book solely on Godwin, N; necrology of Godwin, Q; quoted material from Godwin, W; work dealing in part with Godwin and combinations of them.

The chronological lists of all entries is a mere listing of item number accompanied with these abbreviations. But a cross-tabulation with languages used or place of publication could have shown Godwin's international influence quantitatively. And chronological listing of quotations showing from what material of Godwin one had quoted would have shown what aspects of Godwin had attracted the public attention. This may be a variant of the citation index often used in natural sciences.

Pollin's treatment lacks a taxonomic point of view. Although his first version before publication classified all the material into three groups i) those written in the days when Godwin was working actively, ii) in Victorian period and iii) in modern times, he suppressed all these distinctions into two sections before and after 1836 because of the smallness of the

---

11 See Campanella [B-2] and Pollin [B-5].
amounts published in Victorian period. But the waves of the number of publication are
the reflection of the fashion of thoughts of the day. Relevant classification will reveal the
change of thoughts more clearly. KWIC index applied to the titles of each period might
suggest the difference of thoughts more clearly. For the selection of the title words are
governed by the frame of reference of the original writer. And this frame of reference is
difficult to escape the fashion of the day. Thus the analysis of the difference of the distri-
bution of keywords will supply the quantitative basis for this comparison.22

One might say that these suggestions were beyond the scope of a bibliography. But
such an intensive bibliography as Pollin's one itself provides one interpretation of Godwin's
thought and influence and implications in Pollin's bibliography should be fully visualized
through quantification of the bibliographical contents and abstracts.

We will show this kind of quantification through our experiments on Saint-Simon and
Saint-Simonians' works. Before going further it may be allowed to make some digression
on late Professor Juro Tedzuka whose endeavours, as a result, had provided the basis for
our experiments. Although he is known as one of the most excellent pioneers of mathe-
matical economics in Japan, his concern was not restricted to the mere refinement of
mathematical treatment of economic models but extended to the philosophical implications
underlying the economic theories and the possibility of the social reform from the side of
libertarian favouring P.-J. Proudhon.23 And he spent his stay in Paris from 1921 to 1926
in collecting books and brochures and copying unobtainable ones at Bibliothèque Nationale
for this purpose. His collection covers from Pierre Bayle, Castel de Saint-Pierre, Ency-
clopedistes to contemporary radicals Andre Lorulot and Max Nettlau, especially strong
in the nineteenth century visionaries like Saint-Simon, Fourier, Cabet, Considerant and
their followers. After his death his whole collection was donated to Otaru University of
Commerce by the courtesy of Miyakichi Itaya.

It is on this collection that our experiment on Saint-Simon and Saint-Simonians’
works was undertaken. What he left was not mere books but he also left us one disciple,
Mr. Kiyonari Kidahashi who compiled the first catalogue of Tedzuka collection and de-
signed a hand-made KWOC (Key-Word-Out-of-Context) like index for the section Philo-

---

22 One may easily find this way of investigation as a variant of a quantitative symbol analysis or, more
generally, a content analysis. For example, see Richard L. Merritt, Symbols of American Community,

If we extend our scope beyond the bibliography making, much suggestions will be obtained through
the trials in the field of the computerized information processing found in the journals like Computers
and the Humanities edited by Queens College since 1967.

23 Tamotsu Matsuura describes him as Toshiro Tezuka not as Juro Tedzuka in his “Marginalism in
Japan” in The Marginal Revolution in Economics: Interpretation and evaluation edited by R.D.C. Coats
in his articles like a contribution to W.L. Valk, The Principles of Wages appeared as a Bibliography
annexed by J.Tedzuka.

As to his life and works, see an anonimous article, “Recherche d'économie politique du professeur
Tedzuka et son mérite,” in Bibliographie des Idées sociales et économiques en France; Catalogue de la
bibliothèque Tedzuka de l'université d'Otaru, Redacteur Taro Sakata, Tokyo, 1966.
**sophie et science naturelle** in late forties. His trial found no supporter or collaborator and was isolated until we found a precursor of H.P. Luhn in him. And after twenty years interval this small brochure of about one hundred pages drove us to extend his intention through computerisation.

The importance of Saint-Simon and Saint-Simonians' role in the history of social and economic thought requires no additional explanation. Our data base is composed of the bibliographical description of books and pamphlets written by and on Saint-Simon and Saint-Simonians excluding articles in periodicals. The data base is subdivided into two files; one is for the books written by Saint-Simon and Saint-Simonians themselves (hereafter called BY-SS data base) and the other is for the works written about Saint-Simon and Saint-Simonians' works and their activities (hereafter called ON-SS data base.) BY-SS data base is not complete and we are preparing the second version through international cooperations. Thus the results obtained are rather tentative. ON-SS data base is based on three major information sources; i) 70 items through direct access to the books in Tedzuka Collection and 6 items in Léon Centnell Collection now in Otaru University of Commerce as Ohnishi-Tedzuka Memorial Library, ii) 131 items depending on the description of the secondary information source, Walch's *Bibliographie du Saint-simonisme*, and iii) 30 items from various information sources including our private collections.

Indexes obtained from these data bases are i) KWIC index, ii) chronological index, iii) title index and iv) author index. Other information produced from these data bases is vi) matrix of publication place and date, vii) matrix of language and date and viii) various statistics based on these indexes, etc.

Table 8 is the distribution of publication date obtained from BY-SS ON-SS chronological indexes. This clearly shows the interaction of BY-SS and ON-SS. The upheaval of the sect of Saint-Simonians brought out a burst of publication in 1830s and their revivals are found in BY-SS data three times; at the end of the nineteenth century, after World War I and in 1960s. In most cases the waves of ON-SS data follows BY-SS data. First

---

24 This is an index of 888 headings to 1993 titles for the Section; Philosophie et science naturelle of Tedzuka's collection. An facsimile example is shown below. See *Bibliographie de feu prof. Juro Tedzuka: Section Philosophie. Table analytique*, [rédigée par Kiyonari Kidahashi], x, 100 p. (copie carbonée), [n.d., Otaru],

---

25 This collection was donnated by the alumni to commemorate the sixtieth anniversary of this university. As to the detail of the collection, see Matsuda [13].
TABLE 8. NUMBER OF PUBLICATIONS CLASSIFIED BY DATE

<table>
<thead>
<tr>
<th>Period</th>
<th>BY-SS</th>
<th>ON-SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1802—1820</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1821—1830</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>1831—1840</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td>1841—1850</td>
<td>218</td>
<td>14</td>
</tr>
<tr>
<td>1851—1860</td>
<td>19</td>
<td>10</td>
</tr>
<tr>
<td>1861—1870</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>1871—1880</td>
<td>65</td>
<td>13</td>
</tr>
<tr>
<td>1881—1890</td>
<td>37</td>
<td>10</td>
</tr>
<tr>
<td>1891—1900</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>1901—1910</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>1911—1920</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>1921—1930</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>1931—1940</td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>1941—1950</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>1951—1960</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>1961—1970</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>1971</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

* This figure includes many brochures whose publication dates are assigned to our estimation mainly based on Fournel's Bibliographie saint-simonienne, de 1802 du 31 décembre 1832. 1833.

Upheavals in 1830-50 is the result of the controversies and interests of the contemporaries of Saint-Simonians. The second upheaval at the end of nineteenth century was brought out by the publication of Œuvres de Saint-Simon & d'Enfantin which made possible to utilize many unpublished documents. It should be noticed that the KWIC index shows the increase of SOCIALISME, FEMINISME, POSITIVISME. This may be the change of frame of reference of this period. The second increase in 1920s may relate to Neo-Saint-simonism and the Russian Revolution. Increase of monographs in English in 1940s may reflect the totalitarian political movements of the day. These changes of interests of the time are seen in the KWIC index, and the difference of frame of references between Saint-Simonians and researchers is also clearly observed in the KWIC index. For example, Saint-Simonians preferred the words, ORGANISATION, ASSOCIATION, RELIGION etc. to SOCIALISME, UTOPIE etc.\(^{26}\)

Further application of content analysis requires much more elaborate data base which at least includes additional descriptors or abstracts. Another limitation of this data base is its exclusion of other synchronic social reformers' works, which are necessary to follow their mutual influences. Tedzuka Collection and Ohnishi-Tedzuka Memorial Library are qualified enough to provide materials to compile a data base suitable for this purpose. But as our initial concern was limited to estimate the effectiveness of the automatised indexing and abstracting processes, such kind of elaboration of data base is postponed to the next stage of our experiment.

Some words should be added as to the effectiveness of a retrieval technique of this kind. In the previous section we made a reservation for the final evaluation of classification scheme approach in case of interdisciplinary approach. But in the case of

\(^{26}\) See also Matsuda [13] and Sato's introduction to Matsuda [B-6, no. 3].
retrospective search the superiority of mechanized information retrieval over classification scheme approach is quite clear. Evolution or expansion of classification scheme is itself a result of the division of research disciplines and change of the frames of reference. This means the applicability of older classification scheme to older documents which will be retrieved in retrospective search. Unfortunately change of frames of reference differs from topics to topics. Thus the mechanical conversion or application of the classification scheme has no practical meaning.

III. Association of Keywords and SDI System

Along with KWIC index system, H.P. Luhn had developed the idea of SDI (Selective Dissimination of Information) System where newly arrived data base will be screened at once through a compound list of the descriptors or keywords which each user had prescribed in order to avoid repeated scanning of the same machine readable data many times for each user. The practical applicability of SDI to social sciences might be severely limited because of the lack in massive machine readable data. But even if such data were available, there would remain another factor which will suppress the effectiveness of SDI for the social sciences. For, exact meaning of the categories used in the social sciences often depends on the context and exact specification of descriptors is quite difficult.

The same factor will, in turn, favour the KWIC index. The effectiveness of KWIC index depends on the degree of the context-dependency of the category shown by the keyword. Exact meaning of a keyword will be brought out by the association with other words surrounding the keyword. Thus the effectiveness of KWIC index and similar techniques cannot be determined uniquely. It depends on the user’s insight. This fact is applied to another aspect of a KWIC index retrieval process. The effectiveness will be increased with the association of keywords through glancing the column of keywords.

Some might say that this kind of arbitrariness is a defect of the system. But this arbitrariness should be regarded as the flexibility of the system which can create the room for learning process in retrieving process in action. For the rigid SDI system provides new information aspect only through assigning new descriptors and so learning process at the retrieving process occurs within a limit of user’s memory and documents retrieved previously. On the other hand, glancing the column of keywords will supply a new frame of association of categories which might be overlooked by the researcher in the case of SDI system. Our hypothesis underlying this way of thinking is that it is impossible for the searcher of information to grasp the whole idea of what he wants to retrieve. This might sound ridiculous but is much closer to the reality. Thus the effectiveness of any retrieval technique still rests on the skill or insight of the user.

Another limitation to the information retrieval system, however, comes from the extent of the exhaustiveness of the documents to be retrieved and the degree of additions of descriptors or abstracts which are utilized for retrieval. Like production of goods, no input creates no output but inappropriate input creates no output either. Similarly no

---

57 The association here defined is not the same as to the association found in Lauren B. Doyle’s "Indexing and abstracting by association" in Elias [6] but may correspond to the abduction process of perception.
# APPENDIX I. MAIN FEATURES OF THE

<table>
<thead>
<tr>
<th>Reference no.</th>
<th>1 (B-4)</th>
<th>2 (B-1)</th>
<th>3 (B-5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organizer</td>
<td>Kenneth Janda</td>
<td>Joan Aldous &amp; R. Hill</td>
<td>Burton R. Pollin</td>
</tr>
<tr>
<td>Publication</td>
<td>1964</td>
<td>1967</td>
<td>1967</td>
</tr>
<tr>
<td>Subject field</td>
<td>Political science</td>
<td>Marriage &amp; family</td>
<td>On William Godwin</td>
</tr>
<tr>
<td>Coverage of publication date</td>
<td>1906-1963</td>
<td>1900-1964</td>
<td>1783-1966</td>
</tr>
<tr>
<td>Coverage of publication form; A, articles, B, books.</td>
<td>A</td>
<td>A,B</td>
<td>A,B</td>
</tr>
<tr>
<td>Source materials; D, Direct access or primary sources, S, secondary sources</td>
<td>D</td>
<td>S</td>
<td>D,S</td>
</tr>
<tr>
<td>Selection criteria; M, mechanical, S, selective</td>
<td>M*</td>
<td>S*</td>
<td>S</td>
</tr>
<tr>
<td>Treatment of languages other than English</td>
<td>not applicable</td>
<td>English translation</td>
<td>Original language in Roman alphabet transcription with English translation</td>
</tr>
<tr>
<td>Items covered</td>
<td>2,614</td>
<td>12,850</td>
<td>3,379</td>
</tr>
<tr>
<td>Data base compiler</td>
<td>K. Janda and others</td>
<td>J. Aldous &amp; R. Hill with international corporation</td>
<td>B.R. Pollin with international corporation</td>
</tr>
<tr>
<td>Notes</td>
<td>*All items in the American Political Science Review.</td>
<td>*Selection aims to select all relevant materials based on the various reference books and not to exclude unimportant materials. This procedure is based on the efforts of the Minnesota Inventory of Published Research on Marriage &amp; Family Behavior since 1955.</td>
<td></td>
</tr>
</tbody>
</table>
### Bibliographies Discussed in the Text

<table>
<thead>
<tr>
<th>4 [B-3]</th>
<th>5 (not yet published)</th>
<th>6 [B-6, no. 1]</th>
<th>7 [B-6, no. 33]</th>
<th>8 [B-6, no. 4]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malcom H. Gotterer</td>
<td>Yoshiro Matsuda</td>
<td>Yoshiro Matsuda</td>
<td>Yoshiro Matsuda</td>
<td>Yoshiro Matsuda</td>
</tr>
<tr>
<td>Computer management</td>
<td>Management sciences</td>
<td>Econometrics</td>
<td>Saint-Simon &amp; Saint-Simonians</td>
<td>Library &amp; information science</td>
</tr>
<tr>
<td>A,B</td>
<td>B</td>
<td>B</td>
<td>B*</td>
<td>B</td>
</tr>
<tr>
<td>n.a.</td>
<td>D</td>
<td>S</td>
<td>D,S</td>
<td>D</td>
</tr>
<tr>
<td>S</td>
<td>M*</td>
<td>M*</td>
<td>S</td>
<td>S*</td>
</tr>
<tr>
<td>not applicable</td>
<td>Original languages in Roman alphabet transcription</td>
<td>Original language in Roman alphabet transcription with English translation</td>
<td>Original language in Roman alphabet transcription</td>
<td>Original language in Roman alphabet transcription</td>
</tr>
<tr>
<td>900 above</td>
<td>1,136</td>
<td>864</td>
<td>455 (texts), 231 (criticism)</td>
<td>1785</td>
</tr>
</tbody>
</table>

*All items purchased by the Department of Management Sciences. Selection for purchase was done by teaching staffs of the department.*

*All items reviewed in Econometrica. Editor of this review section was Gerhard Tintner.*

*Books contain some offprints from journals. As to the criticism the articles in journals will be issued.*

*All items held in the public and university libraries in Hokkaido District. Selection was carried out by the cooperation of the Hokkaido Librarians’ Study Circle.*
<table>
<thead>
<tr>
<th>Reference no.</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master list</td>
<td><strong>Author alphabetized bibliography (listing by alphabetized authors)</strong></td>
<td><strong>Complete reference list (listing by alphabetized authors)</strong></td>
<td><strong>Synoptic bibliography (classified by publishing dates before and after 1837, books or periodicals)</strong></td>
</tr>
<tr>
<td>Indexes</td>
<td>Name approach</td>
<td>Author</td>
<td>Author-cross reference (author and coauthor)</td>
</tr>
<tr>
<td></td>
<td>Persons mentioned</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Title</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Titles cited</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Periodical title</td>
<td>+</td>
<td>(*)Periodicals list (not index)</td>
</tr>
<tr>
<td></td>
<td>(Combined with the synoptic bib. listed by alphabetized periodicals)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subject approach</td>
<td>Natural language</td>
<td><strong>Keyword listing (i.e. KWIC index)</strong></td>
<td><strong>KWIC index</strong></td>
</tr>
<tr>
<td></td>
<td>Descriptors</td>
<td>-</td>
<td>Subject index (i.e. classified index)</td>
</tr>
<tr>
<td></td>
<td>Classified scheme</td>
<td>-</td>
<td>*(Publication form)</td>
</tr>
<tr>
<td>Evaluation approach</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Chronological approach</td>
<td>-</td>
<td>-</td>
<td>(Only statistics)</td>
</tr>
<tr>
<td>Language approach</td>
<td>-</td>
<td>-</td>
<td>(Only statistics)</td>
</tr>
<tr>
<td>Number of indexes</td>
<td>3</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>Number of statistics</td>
<td>0</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Computational work</td>
<td>Number of data cards</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Format of masterlist</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Pages (catalogue, illustrations)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>n.a.</td>
<td>B-5, 90 lines/1 column (225,12)</td>
<td>50,000 above A-4,107 lines/2 columns (225,12)</td>
</tr>
<tr>
<td>Computing facility</td>
<td>Northwestern Univ. Computing Center</td>
<td>Univ. of Minnesota Data Processing Center</td>
<td>New York Univ. The Institute for Computer Research in the Humanities</td>
</tr>
<tr>
<td>Computer</td>
<td>Programming language</td>
<td>IBM 709 MAP, FORTRAN</td>
<td>IBM 1401 n.a.</td>
</tr>
<tr>
<td>Programmer</td>
<td>James S. Aagaard</td>
<td>Ralph J. Willard, Ernst Duffing</td>
<td>George W. Logmann</td>
</tr>
</tbody>
</table>

Notes:
* Index indicating the item number.
** Index combined with short bibliographical description.
- Not tried.
+ Not applicable.
<table>
<thead>
<tr>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bibliography</strong></td>
<td><strong>Master list (classified)</strong></td>
<td><strong>Master list</strong></td>
<td><strong>Master list</strong></td>
<td><strong>Main bibliography; classified</strong></td>
</tr>
<tr>
<td><em>Author index (author, co-author)</em></td>
<td><em>Author index (author, co-author)</em></td>
<td><em>Author index (author, co-author)</em></td>
<td><em>Author index (author, co-author)</em></td>
<td><em>Author index (author, corporate author)</em></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td><strong>Title index</strong></td>
<td><strong>Title index</strong></td>
<td><strong>Title index</strong></td>
<td><strong>Title index</strong></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td><strong>KWIC index</strong></td>
<td><strong>KWIC index</strong></td>
<td><strong>KWIC index</strong></td>
<td><strong>KWIC index</strong></td>
<td><strong>KWIC index</strong></td>
</tr>
<tr>
<td>-</td>
<td><strong>LCSH index (subject headings)</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td><strong>(Combined with master list)</strong></td>
<td>-</td>
<td><strong>(Texts; publication form), (Criticism)</strong></td>
<td><strong>(Combined with main bibliography)</strong></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td><strong>Chronological index</strong></td>
<td><strong>Chronological index</strong></td>
<td><strong>Chronological index</strong></td>
<td><strong>Source book reference</strong></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><em>Key to the languages other than English</em></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td><em>Key to languages other than English</em></td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>56</td>
</tr>
<tr>
<td>n.a. B-5, 60 lines/1 column (152,2)</td>
<td>n.a. A-4, 100 lines/2 column</td>
<td>A-4, 100 lines/2 columns (145,32)</td>
<td>A-4, 100 lines/2 columns (92,32; texts), (49,32 criticism)</td>
<td>A-4, 104 lines/2 columns (392, 19, and statistics)</td>
</tr>
<tr>
<td>Johns Hopkins Univ.</td>
<td>Hokkaido University Computing Center</td>
<td>Hokkaido University Computing Center</td>
<td>Hokkaido Univ. Computing Center &amp; Kodak Micro-System Service Center</td>
<td></td>
</tr>
</tbody>
</table>
information will be obtained by inappropriate data base. Thus the real problem is how to compile data base for a bibliography or catalogue.

Apart from this problem of adequacy, the exhaustiveness or thoroughness of coverage requires a close contact with various institutions in international scope in the case of the social science bibliography which, more or less, contains some historical and international aspects. This requirement will be fulfilled by the process similar to compiling a subject union catalogue. Unlike the national or aggregated union catalogue, subject union catalogue requires selection from the whole holdings of institutions participating in the union catalogue making. In this case, not a card catalogue but a check list equipped with many indexes will decrease the burden of librarians to corporate with the program, because a book form catalogue is easy to scan. Thus, not a mere listing of the items to be searched but a bibliography fully equipped with indexes like a final draft is necessary as a check list for the holding survey. For this purpose, the machine readable cataloguing technique is the most powerful weapon for editing. For, once several indexes are prepared, often changes of the item numbers causes serious difficulty in the case of hand made catalogue. So in most cases tentative drafts are mere listing of whole items to avoid the overall re-numbering which will be caused by the adjustment of the item numbers for the new additions or deleting. But, by virtue of the computer processing, this kind of adjustment comes to be a fairly easy task and so the circulation of the drafts of bibliography is possible for any version of the drafts. While it is not quite easy for computer processing to prepare the drafts which can be circulated several times with ease, recent development of COM (computer output microfilm) device provides a good means of producing a circulating version.

Thus our final remark is that the future of the computerised bibliography making depends on the piling up of the trials of compilation of such bibliography in practice. We hope this short essay will be a stimulus for the future trials.

APPENDIX 2. DATA BASE AND SOME TECHNICAL NOTES

The data base for [B-6] was first restricted by the compiler of PL / 1 of FACOM 230-60. The area of characters is within 256 letters, and so the title of the document is restricted within 256 letters including blanks between words. The input format is fixes field and one record is 1,000 letters in case of [B-6, no. 1-no. 3]. Extension to 1,400 letters is executed for [no. 4] to utilize note area.

The master file format is shown in the Figure 2, below. As this is a final print out format, the end of the line is adjusted by the hair space technique. Fig. 3 shows the example from [B-6, no 4] which doubly prints out the titles having articles at the top as omitting article case and including article case. As to the detail of the programming techniques, see Matsui [16].
FIG. 2. DATA BASE
Management Sciences (not yet published, Ref. no. 5 in Appendix 1)

507 E
K 8- 1- 70

TARSKI, ALFRED
INTRODUCTION TO LOGIC AND TO THE METHODOLOGY OF
DEDUCTIVE SCIENCES
1941 NEW YORK
LCIH1 MATHEMATICS--PHILOSOPHY
LCIH2 ARITHMETIC--FOUNDATIONS

[B-6, no. 1]

2333 J
OHOKAWA, KAZUSHI
NGOYO NO DOTAI BUNSEKI (DYNAMIC ANALYSIS OF
AGRICULTURE)
1958 TOKYO: JOSUI SHOBON
(=MFTR 27- 3- 497 REV. BY NAKAMUKA-JAMES I.)

2334 F
ALLAIS, MAURICE
FONDEMENTS D'UNE THEORIE POSITIVE D'UNE COMPORTANT
UN RISQUE ET CRITIQUE DES RESULTATS DE L'ECONOMIE AMERICaine
1955 355 PARIS: IMPRIMERIE NATIONALE
(=MFTR 27- 3- 596 REV. BY SUPPES-PATRICK.)

2335 E
MEADE, J.E.
TRADE AND WELFARE. (WITH MATHEMATIC SUPPLEMENT)
1955 615P LONDON: OXFORD UNIVERSITY PRESS
(=MFTR 27- 3- 500 REV. BY HARBERGER-ARNOL.)

[B-6, no. 4]

0039 J
HOKKAIDO, MUSASHI JOSHI TANKI DAIGAKU, TOSHOKAN,
TOSHOKAN JUKOKAI SHIROYA KIYUKOKU.-- 1972 NEN 3 GATSU
GENSHI--.
1972 SAPPORO: THE LIBRARY. 77P.
LOC: OTARU U.C.M MUSASHI JC

-- 002 ENCYCLOPEDIAS, DICTIONARIES,
GLOSARIES, HANDBOOKS AND MANUALS ---

(ENCYCLOPEDIAS)

0040 C
KENT, ALLEN (ED)
LANCOUR, HAROLD (ED)
NASH, W. Z. (AS ED)
ENCYCLOPEDIA OF LIBRARY AND INFORMATION SCIENCE. VOL.
1: A TO ASSOCIAC.
1968 NEW YORK: MARCEL DEKKER. 12, 676P.

LCIH1 LIBRARY SCIENCE--DICTIONARIES
LCIH2 INFORMATION SCIENCE--DICTIONARIES
LOC: OTARU U.C.M HUC(TIP) HUC(ENG) HUC(ECW)
HUC(AOD) HUC(SAP) KITAMU I.T SAPPORO C.
HOKAUI GU HOKUSEI GU SAPPORO U. MUSASHI JC
SAPPORO JC HOK.PUE.LB UC/KANAGAW SB/BERZER

(C01) ABBREVIATIONS (ELIS V.I P.1-12)
(SCHTEMBERG, VIRGINIA)

(C02) ABSTRACT CLASSIFICATION (ELIS V.I P.12-16)
(DALIY JAY E.)

(C03) ABSTRACTS AND ABSTRACTING (ELIS V.I P.16-30)
(BERNE, CHARLES L.)

(C05) ACADEMIC STATUS OF LIBRARIANS AND INFORMATION
SCIENTISTS (ELIS V.I P.27-12)
(KHURSHID, ANIS +)
Fournel, Henri J.M.,

BIBLIOTHÈQUE SAINT-SIMONIENNE DE 1802 AU 31 DÉCEMBRE 1833. 1300 PARIS.

(1802-1833)

Fournel, Henri J.M.,

DU CHEMIN DE TERRE AU MARSEILLE PAR LA VALLEE DE LA MARNE. PREMIERE PUBLICATION. 1833. 310 PARIS.

(1833)

Fournel, Henri J.M.,

RÉPONSE À UNE BROCHURE INTITULÉE "MÉTIERS SUR LE DROIT HÉRÉDITAIRE, SPÉCIALEMENT CONSIDÉRÉE DANS LA FAMILLE ET LA MORT; PAR LE BARON MASSIAS." (EXTRAITS DU "GLOIRE"). 1831. 350 PARIS.

(1831)

Fig. 3. Title Index

1469
La passion en France: Genèse et évolution de ses fonctions psychosociales. (Univer...

1490
La passion moderne. (Cinquante ans.)

1492
A laboratory course in reading and writing.

1494
Language and information: Selected essays on their theory and application. (Adams...

1496
The language of the foreign book trade, abbreviations, terms, phrases. 2, ed.

1500
The language of public library service in Canada, with particular reference to the...

1504

1506
Le livre: Histoire, fabrication, achat, classement, usage et entretien.

1508
Le livre: Histoire, fabrication, achat, classement, usage et entretien.

1510
Learning from reference books. (Commonwealth and International Library.

1512
Learnin for libraries.

1514
Lehrbuch der bibliothekswaltung.

1516
Learning library methods.

1518
Les bibliothèques. (Quatre ans.)

1520
Les catalogues du département des imprimes de la bibliothèque nationale.

1522
Les catalogues du département des imprimes de la bibliothèque nationale.

1524
Les insectes ennemis des livres dans le monde, moyens de les exterminer.

1526
Les transformations de la bibliothèque nationale et le dépôt annexe de Versailles.

1528
Lettres sur la bibliothèque.

1530
Lettres sur la bibliothèque.

1532
Lexicon der Buchweisen, 4. ed.

1534
Lexicon der Buchweisen, 4. ed.

1536
Lexicon der Buchweisen, 4. ed.

1538
Lexicon der Buchweisen, 4. ed.

1540
Lexicon der Buchweisen, 4. ed.

1542
Lexicon der Buchweisen, 4. ed.

1544
Lexicon der Buchweisen, 4. ed.

1546
Lexicon der Buchweisen, 4. ed.

1548
Lexicon der Buchweisen, 4. ed.

1550
Lexicon der Buchweisen, 4. ed.

1552
Lexicon der Buchweisen, 4. ed.

1554
Lexicon der Buchweisen, 4. ed.

1556
Lexicon der Buchweisen, 4. ed.

1558
Lexicon der Buchweisen, 4. ed.

1560
Lexicon der Buchweisen, 4. ed.

1562
Lexicon der Buchweisen, 4. ed.

1564
Lexicon der Buchweisen, 4. ed.

1566
Lexicon der Buchweisen, 4. ed.

1568
Lexicon der Buchweisen, 4. ed.

1570
Lexicon der Buchweisen, 4. ed.

D'ainville, Virgin

Deanery, Paul

Dewey, Catherine, A. (J.)

Bar-Hillel, Yehoshua

Cox, Jber

Counez, Violet L.

Lawrence, Edith V.

Paris, bibliothèque nationale

Ch., Albert

Egan, Peter

Hilje, Edward Wyman (J.)

Marber, Wilhelm

Marr, Leo Paul

Mason, Angel

Paris, bibliothèque nationale

Housten, Constant W.

Cain, Julian

Ratt, Guy A.

Kraemer, James (Ed.)

Lofthus, Karl (Ed.)

Wasserman, Paul

Johnson, Byron Lamar (Ed.)

Heller, Frieda Marie

Marvin, Eleonora M.

Marvin, Eleonora M.

Sankey, Wilfred L. (Ed.)

[3000 F]

[B-6, no. 3]
REFERENCES

I. Bibliography cited.


No. 2. *Bibliography; Works by Saint-Simon and Saint-Simonians.* (Introduction by Akira Hirota and others) (to be published)


No. 4. *Bibliography; Works on Library and Information Science, based on the holdings of the libraries in Hokkaido region.* Compiled by RINET (Regional Information Network) Society with the joint auspices of Hokkaido Librarians' Study Circle. Otaru, 1974. 2 vols.

II. Books and articles cited.


