DOCUMENT RESUME

ED 045 104 LT 002 245

AUTHOR Standera, Oldrich
TITLE COMPENDEX Profile Adjustment Manual.
INSTITUTION Information Retrieval Research Council of Alberta, Edmonton.
PUB DATE Nov 70
NOTE 28p.

EDRS PRICE EDRS Price MF-$0.25 HC-$1.30
DESCRIPTORS *Evaluation Techniques, Feedback, *Information Science, *Information Systems, Manuals, *Relevance (Information Retrieval), *Use Studies
IDENTIFIERS Canada, COMPENDEX, *Computerized Engineering Index

ABSTRACT

If an information system is to survive, the users must be satisfied that it meets their needs promptly and consistently. It is essential to react quickly to any undesired result such as an extremely high or low output, too low a relevance or recall, or both. The search editor should feel responsible not only for the profile setup but also for its continual adjustment. To do this he must know: (1) the user's exact information need, (2) the user's relevance and recall and (3) the means to achieve the desired objectives. The COMPENDEX Profiling Guide (LT 002 117) provides guidelines for formulating profiles while this manual is designed for search editors and users who wish to adjust their profiles. The manual contains information concerning the user's exact information need, profile performance and user's preference, and the means to bring about the desired objectives such as how to: (1) increase relevance, (2) improve relevance and (3) increase recall. Additional information concerning evaluation of COMPENDEX (Ei and service of Engineering Index, Inc.) is contained in LT 002 116. (MM)

ALBERTA INFORMATION RETRIEVAL ASSOCIATION

COMPENDEX PROFILE ADJUSTMENT MANUAL

by

OLDRICH STANDERA

INFORMATION SYSTEMS
THE UNIVERSITY OF CALGARY

Issued by

Information Retrieval
Research Council of Alberta
Edmonton 7, Alberta
Canada

November, 1970

CONTENTS

ILLUSTRATIONS

I INTRODUCTION

For any system to survive, the users must be satisfied that it meets their needs promptly and consistently.

Especially in the initial period of testing a user's profile, it is essential to react quickly to any undesired result such as an extremely high or low output, too low a relevance or recall, or both. Otherwise the user will be discouraged and the service discredited. If we are not to oversell the system, we must maintain a good standard quality of service, or the disappointed user will not pay for it.

It is clear that the search editor should feel himself responsible not only for the profile setup but also for its continual adjustment.

If he is to be able to take the proper steps, he must know both what to do and how to do it. His knowledge must cover three areas:

1. The user's exact information need.
2. The user's relevance and recall, and which, if either, he prefers to improve.
3. The means to achieve the desired objectives.

Let us follow these points, making use of the information and "know-how" compiled in our previous work (1).

II PURPOSE

Our COMPENDEX Profiling Guide provided guidelines for formulating profiles. Now, with the growing number of users, the need has arisen for a COMPENDEX Profile Adjustment Manual.

This aid is designed mainly for search editors and for a limited number of users who wish to adjust their own profiles.

Note: COMPENDEX in this manual is the designation of the computerized version of the Engineering Index Monthly, supplied by the Engineering Index Inc., 345 East 47th Street, New York, N. Y.

TEXT-PAC, as referred to in this manual, is IBM's system whose main author is Dr. Samuel Kaufman of IBM.

We use this system to process the Engineering Index tapes for both current awareness and Retrospective Search.

1.1 THE USER'S EXACT INFORMATION NEED

We must always ensure that the user's profile is congruent with his actual information need. While this might seem to be a self-evident postulate, it is not so easily achieved.

The various situations that can occur may be illustrated as follows:

N = the Information Need

Q = the Question

1. The question should cover perfectly the information need. The user gets the pertinent information.

2. The question is formulated more broadly than the actual information need; i.e., Q is larger than N. The user will thus get irrelevant information for the area of Q in excess of N. Relevance is low.

3. The question constructed does not cover all of the information need; i.e., Q is smaller than N. The user will be missing the

information for the area of N in excess of Q, or, recall is too low.

4. In the practical information-retrieval environment, we often meet a situation like this:

Here, the information need and the question partially overlap.

This implies that the user will obtain pertinent hits only for the area of overlap. The remaining area of Q represents the amount of irrelevant information retrieved; the remaining area of N, the relevant information not retrieved.

5. The extreme unfavourable condition is shown in the configuration below, where Q and N fail to completely coincide in any particular.

N Q

It is obvious that (1) represents the most desirable situation. There is, however, another condition which must be met: both N and Q must fall within the area of so-called "*a priori* relevance 'or' mega-relevance."

If the data base is not capable of answering a question satisfactorily, we say it has a low *a priori* relevance.

For example, COMPENDEX data base would not be pertinent to any question of a historical or philosophical nature. In many cases it will be partly suitable.

A priori relevance is the first thing to consider before going any further in the profile setup. The coverage of the data base (subject areas covered, documents abstracted) will provide an answer here.

Once a profile has been set up, both user and search editor are responsible for its continued adjustment in the light of previous runs. All users should report immediately any change in their information need. When a user marks hits as irrelevant, although the information was correctly selected as a match for his profile, it may be that he has changed his subject area of interest and failed to let us know. In such case the user should be asked immediately to either revise or rewrite his profile.

IV PROFILE PERFORMANCE AND USER'S PREFERENCE

In COMPENDEX service, by using TEXT-PAC we are able to increase either relevance at the expense of recall or recall to the detriment of relevance, or to operate with reasonable degrees of both. Our COMPENDEX/TEXT-PAC-CIS Project Report (1) will provide any desired details concerning relevance and recall evaluation, system performance, etc.

Relevance is a very useful indicator of the system performance. There is, however, a big difference between a relevance of 60% that represents 6 out of 10 abstracts, and a relevance of 60% that represents 600 out of 1000. Where 4 irrelevant abstracts are negligible, 400 become a massive inconvenience. The specific amount of "trash" (irrelevant retrieved) must be indicated, then, to make the relevance figure really meaningful.

Recall, though a natural counterpart of relevance, may be rather difficult to determine. If ascertained, it is very appropriate for it to be accompanied by the number of "misses" (relevant not retrieved).

It is an inherent property of relevance and recall that, once a certain standard of performance has been reached, we can improve either only at the expense of the other.

We have seen that to truly characterize the performance of a profile we must relate relevance with "trash" and recall with "misses." However, in order to have a complete picture of a profile, we need to know which the user prefers: high relevance, high recall, or medium degrees of both. We can then make a responsible decision about what to do with the profile, and find the most effective steps to attain this goal. We recommend asking the users to supply this information at their earliest convenience so that we can monitor the output of their profiles in the desired way.

If the status of a profile run looks like this (three months' average):

Profile 800116 (Relevance oriented) Rel 30,T70; Rec 70,M5;

we have to improve relevance.

On the other hand,

Profile 800108 (Recall oriented) Rel 85,T16; Rec 20,M240;

calls for a radical improvement in recall.

It is apparent that a medium-oriented user would be content with this situation:

Profile 800108 (Medium oriented) Rel 59,T13; Rec 52,M30.

Our conclusion should not be based on a random result of a month, but on an average of several months. We will find, however, that if a profile remains unchanged and the coverage in the data base is fairly steady, the output relevance/recall ratio will not vary significantly.

To keep track of the profile performance we suggest that the following format be considered to (Fig. 1):

1970

NAME	PROFILE NO.	PREFERENCE	JAN		FEB		MAR		APR		MAY ...		NOTE
SMITH J. F.	800108	REL	75	15									
			40	60									

Fig. 1 • Profile Performance Form

This record means that our user, SMITH J. F., profile number 800108, indicated relevance as the preferred parameter in viewing his output. In

January 1970 his profile had attained relevance of 75% and there were 15 irrelevant retrieved abstracts (trash). Recall was found to be approximately 40%, which means that some 60 relevant abstracts were not retrieved. Each month's record has the following format:

Rel	Trash
Rec	Miss

The degree of relevance will be established easily. In the profile 800108 above, we have received an output of 60 cards in the January run. We sent this output to the user and we got his feedback. He marked 45 response cards "relevant" by pushing out the appropriate box. The remaining 15 response cards represent the irrelevant portion of the output. As 45 relevant abstracts make 75% of the total output of 60, we say the relevance equals 75% and we note this figure in the upper left-hand corner of his profile in the January column. As 15 abstracts retrieved were irrelevant (trash), we write this figure in the upper right-hand corner.

Recall is more difficult to determine. To find out the relevant information not retrieved, we have to go through the entire data base, unless we use some of the less clumsy methods which, of course, are also less precise (see Project Report COMPENDEX/TEXT-PAC-CIS (1)). We can save a lot of time if we compute recall only for the recall-oriented users. Additional effort can be saved if we determine recall figures for the extreme values of relevance (both high and low), where we may anticipate the extreme opposite values for recall (highest for low relevance and lowest for high relevance); approximate recall values for other profiles can then be derived from the graph drawn between the extremities.

If, for whatever reason, we are unable to establish recall values even in the time-saving ways described, we can still improve recall for the recall-oriented user: we simply apply such of the measures listed in Section 2 of Chapter V as are most appropriate. If even then we do not succeed in obtaining a better recall, the data base may be regarded as

unsuitable for the information request. This, of course, should have been ascertained before submitting the profile.

If we proceed in the time-saving way described above (i.e. we use the method of recall evaluation set forth in (1), and take only the recall-oriented profiles, and consider only extreme values of relevance plus some medium) we can plot a graph giving the approximate performance of the system. By means of this graph we can roughly estimate other recall values from the assessed relevance values, and vice versa. Supposing we have established the performance as depicted in the Fig. 2, we could roughly derive, that the recall values ranging from 17% to 34% might be anticipated for the relevance of 90%. With the recall value of 34% we would be allowed to operate in the range of 32-90% relevance.

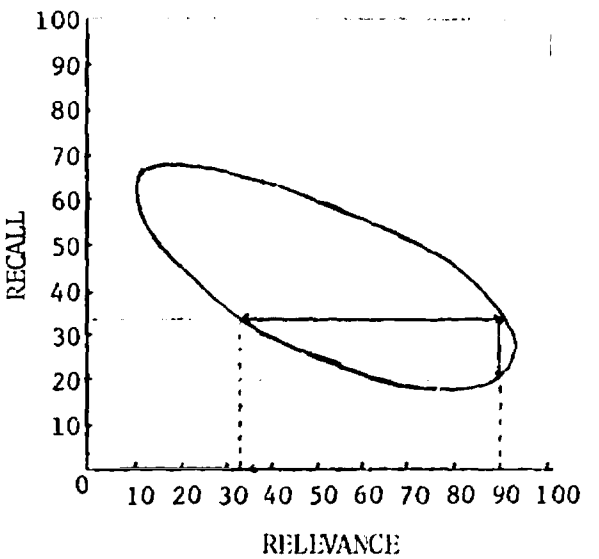

Fig. 2 - System Performance Graph

It remains to clarify the relationship between, on the one hand, high and low output for a profile, and on the other hand, relevance and recall.

High output may be indicative of a higher recall. Usually, some general terms and a very simple structure of the search expressions are responsible

for this type of result. Unless these terms were deliberately chosen, low relevance may be expected in this case.

A low output may be caused either by a low *a priori* relevance of the data base (unsuitable data base for this profile) or by faulty question formulation. In the first case both relevance and recall may be adequate; in the second case, no definitive statement can be made before evaluating relevance, but low recall is likely.

V MEANS TO BRING ABOUT THE DESIRED OBJECTIVES

We should now have reached the point where we know the user's exact information need, the performance of his profile and his preference for relevance or recall.

Let us now examine the ways in which we can optimize system performance and thereby improve user satisfaction.

(1) How to Increase Relevance

There may be some situations contributing to a low relevance which we cannot influence. For example, we may find an abstract which appears to be pertinent, but the hard copy will prove to be irrelevant. The abstractor is responsible for this type of error. He is also responsible if the terms used in a data-base record are too general.

Here are the things the search editor can do to increase relevance:

1. Increase the exhaustivity of the query. ("Exhaustivity" means full, complete, comprehensive definition of an information need.)

Suppose that a user needs information regarding machinery for cutting plastic foams, especially dimensions, prices, production rates, power consumption, working width, automation features. If he does not specify all facets (categories) such as material (plastic, synthetic,

polyurethane), machinery (cutting machine), product (foam), additional features (automation), he may get rather poor relevance.

The search editor should ask the user to add any missing categories and words, to ensure what is called a "multiple approach."

It should be noted, however, that enhancing the exhaustiveness of a profile to promote relevance must be done inside one or more search expressions and not by adding more search expressions. The reason is that search expression (CON) is the entity being searched as a unit.

Remember, expressing the multiple approach by splitting the additional terms into self-sufficient search expressions will not provide a smaller amount of more relevant information, but a higher output with an improved recall.

In formulating a modified profile, a two- or three-level structure should be applied to make a profile more exhaustive (for three-level logic structure see (1)).

2. Increase the specificity of the query. "Specificity" in this sense means a precise definition of a subject using the proper hierarchical level(s). The following example will clarify what is meant. Suppose we want information regarding "thermoplastic resins." The hierarchical structure is (Fig. 3):

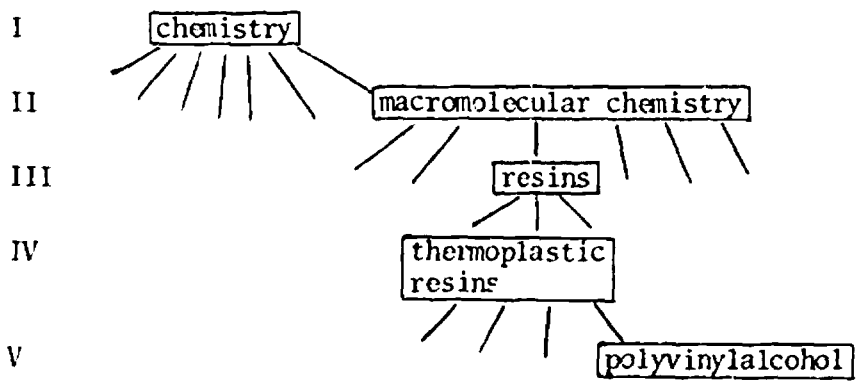

Fig. 3 - Hierarchical Structure

It is evident that specifying the information need as "resins" brings too much noise into the output, whereas stating it as "polyvinyl-alcohol" would prevent other thermoplastic resins being retrieved for the user.

When in doubt, state all hierarchical levels (resins, thermoplastic resins, polyvinylalcohol, etc.) if it is important not to reduce recall.

3. Use "WITH" or "ADJ" where "AND" has produced a low relevance.

Do not use "WITH" and "ADJ" indiscriminately. Use "ADJ" only if there is a fair chance of both words occurring close together in the order indicated; otherwise, use "WITH."

Be sure to take this step only if the user clearly favours higher relevance and does not mind the lower recall that will result.

4. Check whether some irrelevant information could not be traced back to the lack of some "capitalization" feature.

Perhaps incorporating one or two "at" signs in the profile would forestall the selection of some irrelevant documents as hits. If information concerning any acronym is wanted, be sure to use the appropriate capitalization, if the word is a common one in English.

5. "CONTROL" logic can sometimes be helpful in improving relevance. An example of this might be limiting the search to the "author" print control, where "brown" or "smith" would cause irrelevant information to be retrieved in the abstract. Capitalization could do an equivalent job in this case.

6. Consider using "NOT-CONTROL," which may, in a limited number of search expressions, promote relevance.

If, for example, the user is interested in "mining," the search should not be conducted in the "citation" print control, as it would find

any article in the journal "Coal and Mining," and others.

7. Check any truncation in the profile. If truncation is done in a way that leaves the door open for irrelevant items, correct truncation.

Remember, truncation must be tailored to measure. If too broad, it decreases relevance (e.g. "bio" adds "biochemistry," "biography" to the wanted "biology"); if too narrow it decreases recall.

8. Be careful in using a higher match criterion since this may drastically reduce output. The relevance would be improved only if the search expressions, which must be jointly matched, are related. Sometimes profiles are designed so that this indeed will be the case. Otherwise, increasing the match criterion is rather pointless and should be avoided, as it impairs recall. Three levels of vertical structure should suffice to identify the search expression; in rare cases where this is not so, a higher match criterion would be warranted.

9. If an incorrect word has produced irrelevant information, consult the user or another expert in his subject area. This actually should not happen, because correct terms (along with synonyms, antonyms, acronyms and related terms) are the minimum we require from the users. Technical encyclopaedias, thesauri, etc., might also be useful (see COMPENDEX Profiling Guide). Check also for possible errors in coding, typing and punching.

10. Make sure the terms state explicitly the concept desired. Relevance failures are sometimes caused by ambiguous terms.

For example the term "reduction," or any truncated form of it, may supply information dealing with "reduction" as a common type of chemical reaction, but it will also include papers about cost reduction, price reduction, reduction valve, reduction gear, reduction piece, reduced scale, reducing transformer, etc.

So, if the answers are to be directed to cost reduction, specify

"COST$ WITH REDUC$$$$."

11. Make sure there is no false coordination.

For example, suppose we want information about information retrieval and about information specialists using computers. If we formulate our question like this:

A1 RETRIEVAL OR SPECIALIST$
A2 INFORMATION OR COMPUTER$*
CON1 A1 WITH A2

we will be getting not only the desired information, but also the undesired "computer specialist," owing to the false coordination.

Such a construction must be restructured, e.g.:

A1 INFORMATION WITH RETRIEVAL OR SPECIALIST$
A2 COMPUTER$*
CON1 A1 AND A2

12. If a certain formulation of any search expression produces repeated relevance failures, and no part of it can be left out, use the logic "NOT," e.g.:

CON8 NETWORK$.....
CON9 NOT ELECTR$$$$$.....

(2) An Example of Improving Relevance

The information need has been stated as follows: "I request an SDI service in the field of wire production, especially casting. I am interested also in the economical aspects of production. Production costs of wire are of interest as well (in any type of wire production)."

The profile was set up as shown below:

A1 WIRE
A2 MANUFACTUR$$$ OR CAS$$$$ OR PRODUC$$$$
OR COST$ OR ECONOM$$$$
CON1 A1 AND A2

The output for this profile was excessive. The user declared himself to be relevance-oriented.

We discussed the user's profile with him, and as a result made the following adjustment:

1. General terms "manufacture" and "production" were supplemented by the more specific "casting," "prerolling," "rolling," "reeling," "winding."

2. Exhaustivity of this search expression was enhanced by "aluminum" to define the type of material in which the user was actually interested.

We found that only modern continuous methods were sought. "Continuous," therefore, further increased the exhaustivity of this search expression.

3. The truncation CAS$$$$ was wrong and was responsible for retrieving irrelevant "CASE" and "CASING." It was corrected to CAST$$$.

4. Some irrelevant information was a result of the marketing requirement. Because of this we excluded "market," "sales," "prices," "trade," by means of the NOT logic.

5. We tied the logic levels formed by using the more exacting and restrictive WITH in place of AND.

The adjusted profile is designed to yield a high relevance:

A1 WIRE
A2 ALUMINUM OR ALUMINIUM

A3 MANUFACTUR$$$ OR PRODUC$$$$ OR CAST$$$ OR PREROLL$$$
OR ROLL$$$ OR REEL$$$ OR WIND$$$ OR ECONOM$$$$
A4 COST$
A5 CONTINUOUS
CON1 A1 WITH A2 WITH A3 WITH A5
CON2 A1 WITH A3 WITH A4
CON3 NOT MARKET$$$ OR SEL$$$$ OR SALE$ OR PRIC$$$
OR TRADE

(3) How to Increase Recall

1. One very common reason for finding little response in the data base to one's query is lack of multiple approach in the data base. The abstractor or indexer is at fault here. It simply means that the data-base record(s) have few "access points," so that the query, however well-framed it may be, cannot hit the relevant abstracts easily.

These low-quality abstracts will be discovered only by a thorough check of the data base.

The only thing a search editor can do about this, is to formulate the query using all facets (categories); the search expressions should be very simple to increase the chances of any of them hitting into this imperfect description of document content. In other words, we apply higher exhaustivity by using more search expressions (and match criterion equal 1, of course).

Although these tactics are especially suitable for a data base of a low exhaustivity, it will prove helpful any time there is a need to increase recall.

At the same time, the existing search expressions should be made less complicated, if possible.

2. Another way to improve recall is to increase synonymity in a question.

Suppose, for example, we are seeking literature dealing with "search editors," and we are not content with the recall; we specify

INFORMATION WITH OFFICER$ OR SPECIALIST$ OR SCIENTIST$

along with

SEARCH ADJ EDIT$$$

The user should supply all of the synonyms. A user demanding information regarding "potash" can ensure a good recall only if he provides also the terms "potassium carbonate" and "K_2CO_3."

3. A low recall may result if we specify a subject by a term in only one hierarchical level, whereas if specified in two or more levels hits would be very likely.

Referring to our example in the previous chapter, if our information need can be described as "thermoplastic resins," we are more likely to obtain a better recall by using the immediately higher and lower hierarchical levels as well:

RESIN$ OR POLYVINYLALCOHOL

Notice that there is no need to specify "THERMOPLASTIC WITH RESINS," because "RESIN$" covers them too.

Using this type of specification yields lower relevance because of the more general term "resins," which will find other types of resins, too.

4. Recall can be improved by simply using WITH instead of ADJ, AND instead of WITH, or use AND only.

If two words in the question do not match a pair of words close together, they have a better chance of matching them in a sentence or in the whole record. This, of course, does not apply to word groupings which occur only in one particular order (sequence).

5. Great care must be taken in using capitalization. Unless capitalization offers considerable benefit in terms of relevance, it is better to forget it.

It is especially important to remember that acronyms may appear with only initial capitalization as well as with all letters in upper case. If in doubt, it is better to use one "at" sign rather than two.

6. The "CONTROL" feature will very clearly curb output for any profile and decrease recall. Therefore, do not use "CONTROL" if it is desired to enhance recall, unless it is required for some other good reason.

7. "NOT-CONTROL" should be used judiciously, without preventing relevant information from being retrieved.

8. It is important to ensure that any truncation used does not exclude any relevant information. Remember that "ORGANIZ$$$" covers "organize," "organizing," but not "organization," "organizations," "organizational." Think of all forms possible. If in doubt, use unconditional truncation.

9. If higher recall is sought, no match criterion greater than 1 should be used.

10. Words in the profile should be correct as they are supplied by the user. Should it happen by some accident, that they are not, both relevance and recall will be poor. Urge the user to replace them by proper terms.

Misspellings, errors in coding, typing and punching also belong to this category, because they produce words that cannot be matched in the desired way.

Check the entire wording of the profile print-out carefully.

11. Any "NOT" logic used must not be excessively restrictive. For example, in

CON14 CARS OR VEHICLES.....
CON15 NOT MARKET$$$

this "NOT" would be too restrictive if information about prices of cars is also required.

In most cases the "ABS" logic is very helpful, if it is necessary to override any limitation. Suppose information is wanted about agricultural products, but there is enough about barley, rye and oats. If the formulation

CON1 AGRICULTUR$$ WITH PRODUC$$$
CON2 NOT BARLEY OR RYE OR OATS

does not give a good recall for wheat, specify

CON1 AGRICULTUR$$ AND PRODUC$$$
CON2 NOT BARLEY OR RYE OR OATS
CON3 ABS WHEAT

In some cases we cannot help getting a lot of irrelevant material if we want to ensure that certain specific information is picked up. This holds true with general terms such as "profile," "question," etc.

(4) An Example of Increasing Recall

Assume we have the following request: "I need, for a research project I am working on, any information pertaining to the continuous dyeing of synthetic fibers at high temperatures without pressure. I am especially, though not exclusively, interested in the dyeing of polyester fibers. I have heard that the company SYNDYE has been marketing for this purpose some special types of dyestuffs with good dyeing properties and this information would be of high interest."

The profile was formulated like this:

MATCH CRITERION = 2
A1 SYNTHETIC AND FIBER
A2 POLYESTER
A3 HIGH AND TEMPERATURE
A4 WITHOUT AND PRESSURE
A5 CONTINUOUS AND DYEING
CON1 A1 AND A2 AND A3 AND A4 AND A5
CON2 SYDNYE WITH POLYESTER WITH DYEING

The user did not get any hits for this profile. As the user is recall-oriented, we had to try hard to satisfy him before losing him as a subscriber to the service. We did the following:

1. Having due regard for what was explained, we first of all set the match criterion = 1.

2. We then talked to the user and learned that "FIBER" may be spelled also as "FIBRE."

3. We found that synonymous and related to "SYNTHETIC FIBER" are "ARTIFICIAL," "CHEMICAL," "PLASTIC," "NEW." "WITHOUT AND PRESSURE" could also be stated as "LOW AND PRESSURE," "NORMAL AND PRESSURE."

4. We decided to use truncation in the terms "FIBER$," "POLYESTER$," "TEMPERATURE$" and to use DYES$$$$$ to cover DYEING, DYE, DYES, DYESTUFF, DYESTUFFS as well.

5. We discovered one misspelling which had occurred in coding the profile: SYNTYE should be used instead of SYDNYE.

6. Instead of using one search expression ("exhaustivity in one expression") we turned to more and simpler search expressions ("exhaustivity in more expressions") to apply a multiple approach.

7. We replaced AND by WITH in "SYNTHETIC AND FIBER" and AND by AND in "CONTINUOUS AND DYEING" in view of the possible word sequence; for

example, we could meet in the data base "FIBERS MADE OF THE NEW SYNTHETIC RAW MATERIALS," which would not be matched by using the initial formulation.

8. The user has rounded out his request stating that not only fibers, but also yarns, fabrics and knitted goods were of interest.

The adjusted profile has the following structure:

```
MATCH CRITERION = 1
A1 SYNTHETIC OR ARTIFICIAL OR CHEMICAL OR NEW OR PLASTIC
   OR POLYESTERS WITH FIBRE$ OR FIBER$ OR YARN$
   OR FABRIC$ OR HOSIERY OR KNIT$$$
A2 HIGH TEMPERATURE
A3 WITHOUT OR LOW OR NORMAL AIR PRESSURE
A4 CONTINUOUS
A5 DYE$$$$$$
CON1 A1 AND A2 AND A5
CON2 A1 AND A3 AND A5
CON3 A1 AND A4 AND A5
A6 SYNDYE
A7 POLYESTER
CON4 A5 AND A6 AND A7
```

This formulation offers a very good recall with a fair relevance. It allows even an extreme recall to be obtained, if we use combinations of only two logical symbols in the search expressions. Relevance, however, would be impaired very substantially in this case.

VI CONCLUSION

This manual is intended to fill the gap between our Profiling Guide (2) and the Progress Report (1). As the logical extension of the former it uses the knowledge gained in compiling the latter.

As we have seen, there are various tools available to the search

editor for monitoring in the desired way the output, he receives for profiles submitted.

For the search editor to take full advantage of these tools, he must not only know them, but be capable of selecting the proper means in each particular case for achieving the desired goals.

The prerequisite, however, is to ascertain the user's real information need, and whether he is recall- or relevance-oriented. The proper ways to do this have been shown.

VII RECOMMENDED LITERATURE

(1) COMPENDEX/TEXT-PAC-CIS Project Report; Information Systems, The University of Calgary. Report No. 6, August 1970. By Oldrich Standera.

(2) COMPENDEX Profiling Guide; Information Systems, The University of Calgary. June 1970. By Oldrich Standera.