Mirja Iivonen
Department of Information Studies, University of Tampere

Selection of Search Terms as a Meeting Place of Different Discourses

Abstract: Selection of search terms is considered a meeting place of different discourses. Discourse refers to the ways of talking and thinking about a certain topic. There exist concurrently different discourses on the same topic. Although the selection of search terms used to be described as a translation process, it could also be seen as a situation where searchers step into different discourses and select search terms from various sources on the basis of their own previous experience. Based on an empirical study, the paper describes the various sources of search terms. They are controlled vocabularies, assumptions about indexing, documents and their titles, clients’ words, databases, and the searchers’ own search experience. Because the searchers have different experience and are working in different types of work environments, they use these sources differently. The paper also outlines a model for understanding the selection of search terms as a meeting place of different discourses.

1. Introduction

The selection of search terms has usually been described as a translation process. We have learnt to think that during the process the concepts of a search request described in the client’s words will be translated into the words which can be used in a certain information system.

According to previous studies we know, however, that the selection of search terms is not only a translation process but also the seeking of search terms from different sources. Saracevic et al. (1991) investigated the sources of search terms and found that only 38% came from the written search request, 23% were generated by the client during the interaction between the client and the searcher and 39% came from sources other than the client (e.g., thesaurus and relevance feedback). In former studies it has been shown also that the searchers consider various alternatives for search terms on the basis of their own previous experience and their own work environments. Fidel (1986, 1991a, 1991b, 1991c) found that the searchers had certain routines and they used rules to support their selection of search terms. She also found that the searchers who worked in different types of search environments (practical, theoretical and general) selected search terms differently.

We can consider the selection of search terms a meeting place of different discourses. Discourse refers to the ways of talking and thinking about a certain topic. The discourse of welfare state refers to the ways of talking and thinking about welfare state. It is necessary to see that the different discourses of the same topic may exist concurrently. It is also necessary to understand that the discourses of one topic may change over time. The discourses have also their social functions; they have been produced in social practice (Abercrombie et al., 1994, 119-120).

When the searchers select the search terms to describe one certain search topic they have to step into several various discourses where this topic may be discussed and taught differently. The searchers have to try to understand the client’s talks and thoughts. They have to take into account the ways in which controlled vocabularies describe and organize the same topic. They have to meet the ways in which the authors of the documents talk and think about the same topic.
topic. In addition, they have to consider the ways in which the topic may be discussed in the practice of indexing and databases. The searchers step into these various discourses with their own experience. Because the searchers' experience differs from the others' experience, they may meet and understand these discourses differently.

In this paper I will first describe the sources of search terms on the basis of my empirical study. I will show how strong the role of different sources is in the selection of search terms. After that I will outline a model for understanding the selection of search terms as a meeting place of different discourses.

2. Sources of Search Terms: an Empirical Study

In my empirical study related to the consistency in the selection of search concepts and search terms I also paid attention to the sources of search terms. A total of 24 professional searchers from three types of search environments and eight students of information studies participated in my study. They selected search terms for a search of the same 12 search requests. They also justified their selections. In the group of academically oriented search environments there were six searchers working in university libraries and one searcher working in an information bureau offering services to academics in the whole country (Finland). In the group of special service-oriented search environments there were, in addition to five searchers working in special libraries, two searchers working in university libraries offering, however, specific services to specific user groups in very narrow subject areas. In the group of public search environments there were eight searchers working in public libraries. None of the professional searchers were from the same place of work. The data of my study and different search environments are described in greater detail elsewhere (Livonen, 1994, Livonen, 1995).

On the basis of the searchers' explanations it was possible to see that the searchers got ideas for search terms from different sources (see Figure 1). The controlled vocabulary had a very important role as a source of search terms. Almost every searcher (30 searchers out of 32 = 93.8%) referred to it when they justified the selection of search terms. The controlled vocabulary is a language and it is in a very concrete way present in the selection of search terms. It offers a frame through which the world can be interpreted and described. As a discourse the controlled vocabulary offers ways to talk about a certain topic. At the same time, however, it restricts other alternatives very clearly and often also very strongly. In a controlled vocabulary there always exists the function of standardization; both the preferred terms and their relationships have been standardized. On the basis of the searchers' own justifications they seem to be very involved in the discourse of the controlled vocabulary.

Another fairly important source for search terms in my study was documents. Twenty three searchers (71.9%) referred to the documents and their titles in their justification of search terms. They thought of the ways in which the authors of the documents have talked about the topics of the search requests. One searcher who selected search terms "welfare" (in Finnish "hyvinvointi") and "welfare state" (in Finnish "hyvinvointivaltio") from a search request where social problems and people's everyday life in a society were discussed justified her selection as follows:

(Search request: Social problems, alienation from society, and structural change in society, and their influence on the control people have over their everyday lives in an affluent society.) "... I thought first that it (welfare) may occur in the title or in
The percentage of searchers who justified their selection of search terms by referring to the following sources of search terms:
1 = controlled vocabulary
2 = documents (e.g., the titles of the books)
3 = practice of indexing (assumptions about indexing)
4 = clients' words
5 = database
6 = own search experience in general

Fig. 1: Sources of search terms

the subtitle of a book, or in this kind of context . . . This is just about the context. This is about the welfare state. If could find an author or a book from that area, I would offer it to the client. " (A searcher working in special service-oriented search environments)

In a certain way it is a challenge for the searchers to explore those other words with which the topic of a search request may be discussed in the literature. It is a real task to find a right context from the literature for the search request.

In my study it was not only the discourse of the controlled vocabulary in which the searchers were heavily involved. In addition they were involved in the practice of indexing. Several searchers (22/32 = 68.8%) mentioned assumptions related to the indexing process when they justified the selection of search terms. They thought how indexers would have used controlled terms in describing certain topics. The practice of indexing produces ways of talking about a certain topic, and these ways may differ from the ways produced by the authors of the documents. There are certain rules, policies and restrictions in the practice of indexing. Therefore it is understandable that the searchers take into account these ways when they select search terms.

The clients, when they describe the search requests with their own words, bring their own discourses to the search process. The clients' words are not necessarily very familiar to the professional searchers. They can, however, trust the clients' ways of talking about the topic and consider the clients' words as search terms. When the clients name a certain concept with
a certain term they tell more than only the topic of a search request. They also tell about the
frame and tradition they come from because about the same topic is discussed with different
words within different traditions. In my data a few searchers \((14/32 = 43.8\%)\) pointed out
that they selected certain words for search terms because they were mentioned in a search
request. One searcher described her confidence in the client's words as follows:

(\text{Search request: The influence that alternative activities and environmental movements
have on social decision-making and the use of power and on solving global environmental
problems}) \text{"This was difficult. I selected environmental movements and alternative
activities. I would not have noticed that alternative activities... I think that if the search
request could have been formulated in another way, so I could not have noticed it... I would have used environmental movements and nature conservation movements."} (A
searcher working in public search environments)

We might assume that the database for which the search terms were selected could have had
a strong position as a frame in the selection of search terms. Surprisingly, it had not. Only
ten searchers \((31.3\%)\) referred to the database when they justified their selection of search
terms. They considered whether a certain search term could be or could not be useful in the
database.

Similarly only 10 searchers \((31.3\%)\) referred to their own previous search experience
when they justified the selection of search terms. However, 24 subjects out of 32 were
professional searchers who had search experience. Of course they had their own experience
although they did not explain it in words. The searchers always act on the basis of their own
experience and they cannot exclude it from their action.

There were interesting differences in the use of different sources in the selection of search
terms between the searchers working in the different types of search environments (see Table
1). Only the controlled vocabulary was referred to as a source of search terms equally and
very strongly in each type of search environment. This tells about the strong position of
controlled vocabularies in the selection of search terms.

The searchers working in academically oriented and special service-oriented search
environments took into account documents and their titles more often than the students or the
searchers working in public search environments. The collections of documents are more
limited in academically oriented and special service-oriented search environments than in public
search environments where all possible subject areas are included. It was therefore easier for
the searchers in these environments to think about the discourse of documents. Possibly the
restricted collections helped the searchers to understand and get a specific picture about the
literature of the subject area of the search requests.

The clients' words were referred to mostly by the searchers working in academically
oriented search environments. The clients in this type of search environment were experts in
their own domain. It is therefore easy to understand why the searchers trusted their clients'
own words, considered them and took them into account as search terms.

The professional searchers referred more often to the indexing process and to their own
search experience than the students. The students did not have yet work experience—neither
in the area of information storage nor in the area of information retrieval—and they could not
use it in the selection of search terms. The searchers' own work experience affects their
decisions and when they lack this experience they cannot use it.

On the basis of the searchers' justifications for the selection of search terms we can find
that the searchers know the sources of possible search terms differently. Therefore, their
### Table 1: Sources of search terms mentioned by the searchers working in the different types of search environments

<table>
<thead>
<tr>
<th>Source of search terms</th>
<th>Type of search environments</th>
<th>All search environments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Special service-oriented search environment</td>
<td>Academically oriented search environment</td>
</tr>
<tr>
<td></td>
<td>(n=9) (%)</td>
<td>(n=7) (%)</td>
</tr>
<tr>
<td>controlled vocabulary documents</td>
<td>9 100.0</td>
<td>6 85.7</td>
</tr>
<tr>
<td>practice of indexing clients' words</td>
<td>8 88.9</td>
<td>6 85.7</td>
</tr>
<tr>
<td>database</td>
<td>7 77.8</td>
<td>6 85.7</td>
</tr>
<tr>
<td>own search experience</td>
<td>2 22.2</td>
<td>5 71.4</td>
</tr>
</tbody>
</table>

n = number of searchers who referred to this source of search terms

abilities to take these sources into account vary. We cannot expect all searchers to behave similarly.

### 3. Meeting Place of Different Discourses

The selection of search terms is a meeting place of different discourses. The searchers step into these different discourses on the basis of their own experience and are therefore differently involved in these discourses. The searchers have to consider the ways of talking and thinking about the same topic inside the various discourses. It is obvious that the topic of the search request it has been talked about differently in the documents than in the controlled vocabulary. The controlled vocabulary may rule out alternatives of talking about the topic. It is also obvious that the practice of indexing may have some certain rules and ways which differ clearly from the practice of clients. Many times it is even difficult for clients to understand in which way and why just in this way the topic is described in indexing.

The searchers are (or at least they should be) conscious of these different discourses. They can think of alternatives for search terms from this point of view. The selection of search terms can be described in a model where the searchers consider the possible search terms stepping into the different discourses (see Figure 2).

### 4. Discussion

The searchers are not satisfied with being only translators when they select search terms. They are more like explorers in different discourses. They consider different alternatives from different sources knowing that about the same topic may be talked of differently in various discourses. They also know that the ways of talking about one topic vary and change considerably across time and place.

Although there are several different sources for search terms all of them are not as important as others and are not equally considered by all searchers. The controlled vocabulary has, however, a very strong position in the selection of search terms in general. Therefore, it is not a question of secondary importance which kind of discourse the controlled vocabulary offers.
Although there are several different sources for search terms all of them are not as important as others and are not equally considered by all searchers. The controlled vocabulary has, however, a very strong position in the selection of search terms in general. Therefore, it is not a question of secondary importance which kind of discourse the controlled vocabulary offers.

The understanding of the selection of search terms as a translation process may lead too easily to the replacement of one word (a client's word) with another (a word which can be used in a system), to replace one word with only another one. The understanding of the selection of search terms as a meeting place of different discourses leads to seek and explore search terms from different sources and to compare them. The fact is that there exist many alternative search terms for one topic and many various words can replace one word named in a search request. The fact also is that the same topic may be talked of differently in different discourses. For the searchers there exists the horn of plenty or a treasure of words to be used.

Notes

1. I would like to thank Dr. Diane Sonnenwald and Dr. Leah Leavrouw for the idea of this model. I got the idea for my model from their article (1996), which I translated into Finnish. They presented a provisional model of communication in the information system design process and showed the participants' interactions in design process in various phases. (See also Sonnenwald, 1993, 1995.)
References


