Designing Usable Keyword Search Systems

Marc Hassenzahl[*] and Jochen Prümper[#]
*Siemens AG, Corporate Technology - User Interface Design, D-81730 Munich
marc.hassenzahl@mchp.siemens.de
#FHTW - Fachhochschule für Technik und Wirtschaft, D-10313 Berlin
j.pruemper@fhtw-berlin.de

1 Introduction

The fast and direct access to information is one main advantage of computer technology. Nowadays, due to the enormous memory capacity of a single personal computer hard disk and the vast amount of information offered by the internet, search systems are inevitable. Innovative ideas of how to overcome the difficulties of formulating search queries (e.g. Graß, Haubner & Arend 1999) and the visualisation of information do exist (e.g. Shneiderman 1996). Nevertheless, most of these ideas are still at an experimental stage or require cutting-edge technology.

We recently suggested some guidelines for the design of keyword search systems based upon more conventional technology (Hassenzahl & Prümper 1999, see table 1). The guidelines were derived from an empirical user survey which focused on generalised search strategies (e.g. „I use synonyms") and boolean operators (AND, OR, NOT), thereby explicitly taking computer and internet experience into account. These guidelines can be characterised as a „snapshot" of user expectations. By using them as „design hypotheses" the usability of the search system should be increased.

The objective of the present paper is to explore the suggested guidelines' practical use for designing the keyword search system of a web-based information system.

2 Search system and interface

We used the guidelines below to design a search system for the web based German Federal Health Information System „IS-GBE". Figure 1 shows the final design solution.

Table 1: Guidelines for the design of keyword search systems

	Guideline
1	A search system should support the four commonly used search strategies without making the user change pages or screens. These strategies are: • Searching with **one word** • Searching with **two or more words at once** • Presenting an **overview** of the structure and contents of the database (especially novices) • Using **wildcards** (especially experts), e.g. interpreting a given word automatically as the beginning of a keyword
2	A search system should **check for spelling errors** or provide an option for **including similar sounding words** (search for homonyms)
3	A search system for the <u>internet</u> should provide an option for **including synonyms** taken from a thesaurus. The system should show the used synonyms.
4	The boolean operator **AND** should always be the **default** setting. The system should communicate this setting to the user in plain English descriptions of the operator's effect (e.g. „all the words" for AND).
5	A search system can do both without the option to combine different boolean operators and a NOT-operator. Particularly the NOT-operator is often used as a substitute for effective filtering functionality when a user has to reformulate a query.

The final design of the search system conforms to the proposed guidelines, except for the recommendation concerning the overview (Guideline 1). It was given its own section. Instead of a spell checking option, an operator for searching similar sounding words (phonetic matching) was included.

These exceptions and additional features emerged from technical constraints, non-negotiable requirements and results from the usability test of an earlier version.

Additional features are:

- Non-default and advanced operators, such as „exact phrase and its synonyms", are hidden by using a drop down list in order to reduce the „first-sight" complexity of the interface (see figure 1, number 1).

- The user has to decide whether she wants to search all material (full text option) or to put a scope on the header/textual description of the documents (see figure 1, number 2). With the scope selected, the system's response time will be significantly reduced.

- If a query yielded no results, an automatically presented textual hint suggests to use the more advanced operators.
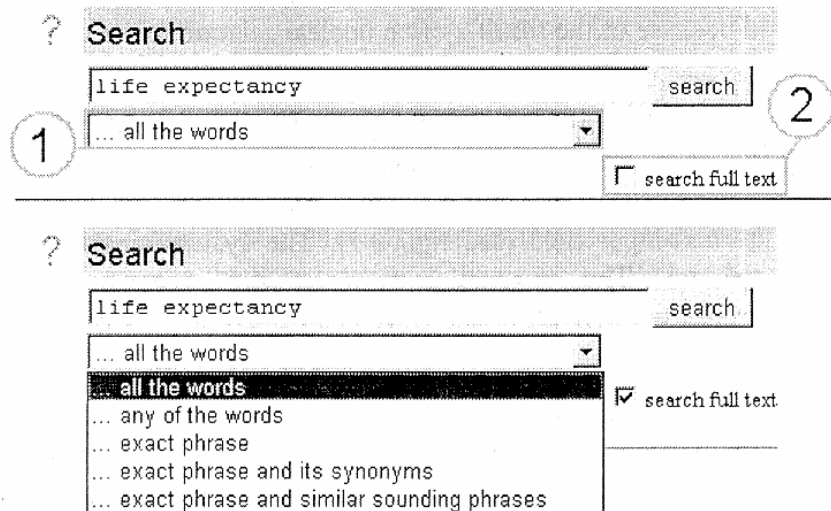

Figure 1: Final design solution of the search system (originally in German)

3 General usability problems

A usability test was performed in order to explore problems of the final design solution. Five subjects participated. All subjects had sufficient work experience and a minimum of two years prior experience with search systems. Two out of five had internet experience. Computer experience was on average.

Subjects worked through seven task scenarios. For instance, participants had to find out the mortality rate for breast cancer in Germany in the year 1982. The scenario required to use the advanced operator „synonyms" to get the intended result, because the term „breast cancer" is not used in official documents. Three of the scenarios focused on the search system alone, three had a wider scope (e.g. searching, collecting and evaluating documents).

All interactions were recorded with Lotus „ScreenCam" and submitted to additional usability problem analysis (Zapf et al. 1992).

This revealed twenty-five different „problem tracks", i.e. participant's behaviour that points to an underlying usability problem, which were further grouped into four general problems.

- *Misspelling/mistyping and difficulties with the „similar sounding words" option*: Misspelling and mistyping proved to be a serious problem theme rising from the lacking spell checking option. Furthermore, there was an uncertainty about whether correct capitalisation plays a role and how to deal

with mutated vowels (e.g. Ernährung - Ernaehrung). The advanced operator „exact phrase and similar sounding words", originally conceived as an alternative to spell checking (see guideline 2), was either not found or not understood because of the result's often unclear relation to the original search word.

- *Including the boolean operators in the query formulation*: More experienced participants connected search words by explicitly typing the required operator, although the operator must be selected with the drop-down list beneath the text-input field (see figure 1). The system simply ignores the typed-in operators. Thus, a situation can arise where users type OR but the system searches with AND. Presumably this behaviour will go unnoticed by the user.

- *Advanced operators are not visible*: The system offered a variety of useful but uncommon advanced operators. If the task scenario explicitly required the use of an advanced operator, participants had problems finding it. This is an effect of the design decision to hide the more advanced operators in order to reduce „first sight" complexity of the interface. Even with textual hints given in cases of a query yielding no results, participants could not overcome this problem. However, this might be an effect of ambiguous wording.

- *Scoped search*: The technology-oriented search scope „full text" versus „header / textual description" was either ignored or not understood.

4 Summary and conclusions

The general problems 1 and 2 point to minor changes and further qualification of the guidelines:

- Guideline 2 proposes spell checking and phonetic matching as alternative options. This cannot be confirmed by our analyses. Spell checking is a basic option which should be included in search systems, whereas phonetic matching is rather rare and possibly hard to understand.

- Guideline 4 has to be further qualified: If there are experienced database users in the intended user group, the explicit input of boolean operators should be enabled. An elegant solution is the automatic detection of the typed operators and a following automatic selection of the appropriate operator from the drop-down list.

Problems 3 and 4 concern aspects mainly connected to the design solution itself:

- The advanced operators incorporated in the present search system make it a powerful tool. Nevertheless, the design decision to hide advanced operators led to most of the functionality going unnoticed by the participants. We

recommend that all possible manipulations of search words should be visible to the user. If that is not possible, active mechanisms have to be employed to make the user familiar with new and uncommon manipulations (advanced operators). Leaving it to the user alone decreases the chance that the system's functionality will be fully used.

- Technology-oriented scoped search is problematic. We propose to have global search of all material as default. If a scope is required, a self-descriptive way of presenting it should be employed. For example, one may frame scoped search as „fast and superficial" and global search as „slow and thorough" to give the user a sense of the scope's effect on his search and its results.

The guidelines as proposed above proved to be helpful for designing the search system studied. As shown above, only minor problems emerged during usability testing, none of them a „usability disaster" (Molich 1994) and all solved easily. Of course, the presented guidelines' application in practice cannot prove their correctness in a scientific sense. However, their application did not cause obvious usability problems. Therefore, the designer may use them for an initial draft in the sense of a first educated guess. A good initial design may reduce the number of iterations necessary to reach a usable final design. This in turn increases efficiency and decreases costs. That is the true power of guidelines.

5 References

Graß, M., Haubner, P. & Arend, U. (1999). "Search Manager": A dynamic and incremental tool for searching complex objects in multihierarchical structures. *Proc of HCI International '99*. In this volume.

Hassenzahl, M. & Prümper, J. (1999). "Benutzererwartung eingebaut": Gestaltungsempfehlungen für Suchfunktionen auf der Basis einer empirischen Benutzerbefragung. In U. Arend, K. Eberleh & K. Pitschke (eds.), *Software-Ergonomie '99: Design von Informationswelten*. Stuttgart: B.G. Teubner.

Molich, R. (1994). Preventing user interface disasters. *Behaviour & Information Technology*, 13(1&2), 154-159.

Shneiderman, B. (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualisation, http://www.isr.umd.edu/TechReports/ISR/1996/TR_96-66/.

Zapf, D., Brodbeck, F.C., Frese, M., Peters, H. & Prümper, J. (1992). Errors in working with office computers: a first validation of a taxonomy for observed errors in a field setting. *International Journal of Human-Computer Interaction*, 4 (4), 311-339.