that we are interested in evaluating in the context of this book. Document search consists of several distinct processes (see Chapter 4), such as translation of documents from natural language into IRL or translation of search requests from natural language into IRL. It is well known that such subprocesses are realized differently in different systems, even if the same IRL is used. Therefore, which realizations of a particular subprocess must be used in a particular situation? To answer this question it is necessary to be able to evaluate these subprocesses. This ability is crucial in designing and developing documentary IR systems, and the subprocesses in question must be included in the list of objects to be evaluated by developers of an IR system. As a matter of fact, we discussed this situation briefly in Chapter 6. Here we will examine the problem in greater detail.

Different approaches are available for evaluating subprocesses. For example, the evaluation may be based on the degree to which a process result corresponds with a standard result (for example, correspondence of the document profile, obtained by the corresponding subprocess, with the standard document profile). In this case, results will be regarded as standard if they have been obtained using the same initial data, while the subprocess has been performed by an expert. However, in this case there is no guarantee that a result (document profile) *close* to the standard one will perform better during the document search than a result that is more *remote* from the standard one. In other words, in this case the evaluation in question cannot predict how the evaluated process will affect document search results. Quite naturally, we are interested in an evaluation that is capable of making such prediction. Therefore we need an alternative approach in forming standard results or an alternative approach in carrying out evaluation.

As for the "correct" approach in forming required standards, to our knowledge it is not yet available in information science. Therefore, let us consider an alternative evaluation approach. This approach to evaluation of a corresponding subprocess is based on the results of searches conducted using this subprocess. Actually, in this case too there is the problem mentioned in Chapter 6, which has no satisfactory solution as yet. The point is that such subprocesses do not form search results alone; instead they form them jointly with other subprocesses available in the system. In such occurrences, there may be a situation in which the considered realization of a certain subprocess in combination with one set of realizations of the rest of the subprocesses will ensure, as a rule, a high-quality search, whereas when combined with another set of realizations of the same subprocesses the search quality, as a rule, will be low. Therefore, when one evaluates such a subprocess using available search results, one must know how other subprocesses have affected these results and take their effects into consideration. As noted, this problem has no satisfactory solution as yet, and regretfully we did not find any published information worth discussing in this book in connection with the problem in question.

It follows, therefore, that we also are not ready to discuss the general case

of the evaluation in question. At the same time, in an important particular case we can present the general outline of the method of evaluation in question. We assume in this particular case that realizations of all subprocesses except the evaluated ones have already been selected and do not change during use of the evaluated realization. The importance of this case is in the fact that it is natural and occurs quite often in practice, for example, in automatic document indexing systems or in systems using "natural" language (i.e., using a document text as the document profile). All search subprocesses in these systems, except constructing query formulations, are realized according to certain algorithms that are unchanged during sufficiently long periods of time. In other words, in such systems during the period of time under consideration different realizations may only be used in the subprocess of constructing query formulations; that is, these very widely used systems belong to the special case that is of interest to us. For brevity we will refer to the subprocess for constructing query formulations as the QF subprocess.

It is clear that in the systems in question searches carried out on the same search request using different realizations of the QF subprocess may lead to different results, these differences resulting solely from the given realizations. Therefore, one may assume that the functional efficiency of a document search in these cases is determined by the used realization; that is, evaluation of functional efficiency of document search in such systems can be regarded as an indirect evaluation of the QF subprocess. If a functional efficiency evaluation is expressed numerically, it will be reasonable to use the determined values as values obtained as an evaluation result for the QF subprocess. For example, an evaluation based on such values will be quite justified if it is used to compare different realizations of the QF subprocess (in the event that the search is conducted on the same search request), because in the context of the discussed evaluation the best realization will ensure the highest functional efficiency of the document search. Note that tasks requiring such a comparative evaluation are of great practical value (see Chapters 6 and 7).

At the same time, the special case discussed earlier may require such an evaluation of the realization of a corresponding subprocess (i.e., a subprocess with a "nonprefixed" realization) that it will permit prediction of results of a search conducted using this realization on a search request that is being newly added to the system. For example, it may be necessary to form a set of the "best" realizations of the subprocess in order to use this set for choosing the most appropriate realization for each incoming search request; or, after forming such a set, it may be necessary to decide whether to include in it a newly proposed realization. There is no doubt that the required evaluation will be more difficult to conduct as compared with the preceding evaluation. However, approaches to make such an evaluation have been proposed in information science. These approaches are mostly based on averaging values of functional efficiency that were obtained in a specially organized series of searches. We will discuss this in detail