

**Принципиальная схема решения задачи
автодополнения пользовательских поисковых
запросов на русском языке и ее анализ**

Голяк Ю.Д.

Белорусский государственный университет, Минск

В работе исследуется проблема автоматического дополнения поисковых запросов для информационных систем с интерактивным пользовательским естественно-языковым интерфейсом.

Цель статьи – постановка задачи автодополнения поискового запроса пользователя, ее анализ, принципиальная схема решения и обоснование требуемой для этого функциональности лингвистического процессора.

Материал и методы. Материалом для исследования послужили массивы реальных поисковых запросов, представленные в открытом доступе в известных базах данных. Методология включает в себя статистический анализ и формально-лингвистические методы.

Результаты и их обсуждение. Функциональность автоматического дополнения пользовательского запроса является неотъемлемой частью поисковых систем. Известные решения задачи, как правило, опираются на автоматически формируемую историю проведенного поиска. Однако существует ряд приложений, для которых такое решение по объективным причинам неприемлемо, чем обусловлена необходимость дальнейших исследований в этой области. В работе предлагается формальная постановка задачи автодополнения запроса и принципиальная схема ее решения с учетом предполагаемой «истории» поиска на базе текстов предметной области. Приводятся классификация типов пользовательских запросов и их синтаксических структур, а также функциональность базового лингвистического процессора, необходимого для решения целевой задачи, и требования к дополнительному модулю формирования подсказок для предполагаемых запросов.

Заключение. Предложенный подход к решению целевой задачи позволяет избежать традиционной зависимости от истории поиска и учесть специфику поискового пространства, обеспечивая, что очень важно, высокорелевантную реакцию поисковой системы.

Ключевые слова: автодополнение, поисковый запрос, предиктивный ввод, автоматический лингвистический анализ текста, синтаксическая структура запроса, лингвистический процессор.

**Conceptual Scheme of the Solution for the Problem
of User Search Query Autocompletion in Russian
and its Analysis**

Goliak J.D.

Belarusian State University, Minsk

The paper investigates the problem of automatic search query completion for information systems with an interactive natural-language user interface.

The purpose of the work is the formal definition of the problem of user search query automatic completion and its analysis, creation of the principle scheme of the solution and description of the functionality of a linguistic processor required for the target task.

Material and methods. The materials for the study include the lists of real search queries available for open access. The research methods are statistical analysis and formal linguistic methods.

Findings and their discussion. The functionality of automatic user search query completion is an essential part of search engines. Common solutions of the problem often rely on an automatically gained user search history. However, there are applications that cannot use this solution for objective reasons, which necessitates further research in this area. The paper proposes the formal definition of the problem of automatic completion of a query and principle scheme of its solution considering the pre-assumed search “history” based on the texts of the domain area. The classification of types of user queries and their syntactic structures, as well as the functionality of the basic linguistic processor required to solve the target problem and the requirements for an additional module for query suggestions generation are described.

Conclusion. The proposed approach to target problem solution gives an opportunity to avoid the traditional dependence on the search history and to take into account the specifics of the search space, providing a highly relevant reaction of the search system.

Key words: Query Type-Ahead, search query, predictive query input, automatic linguistic text analysis, syntactic structure of a query, linguistic processor.