

## ОТЗЫВ

на автореферат диссертации

Голяк Юлии Дмитриевны

*«Предиктивное автодополнение запросов пользователей в системах информационного поиска (на материале русского языка)»*

по специальности 10.02.21 – прикладная и математическая лингвистика

на соискание ученой степени кандидата филологических наук

Судя по теме научного исследования и содержанию автореферата диссертации, работа Ю.Д. Голяк выполнена в русле проблемного поля перспективного лингвистического направления – компьютерной лингвистики. Задачи, решаемые автором, обладают бесспорной научной значимостью и практической ценностью. Сформулированные автором диссертации положения, выносимые на защиту, и выводы, оформленные в результативной части исследования, оригинальны, научно обоснованы и аргументированы, что свидетельствует о высокой профессиональной подготовке соискателя, знании им проблематики исследования и путей решения поставленных задач, а также о значительном вкладе соискателя в науку и практику использования результатов исследования.

Целью диссертационной работы Ю.Д. Голяк является разработка метода, алгоритмов и лингвистического обеспечения решения задачи предиктивного автодополнения русскоязычных запросов пользователей в системах информационного поиска и их реализации в виде промышленного прототипа. Актуальность заявленной темы диссертации и способы ее решения соискателем определяется необходимостью оптимизации существующих поисковых запросов и информационных подсказок с заменой принципа последовательного завершения строки вводимого запроса, рассматриваемого в качестве его префикса, принципом погружения в общий контекст, что существенно повысит релевантность реакции системы по отношению к информационным потребностям пользователей. Основным недостатком существующих решений задачи QTA (этой аббревиатурой соискатель обозначил автодополнение, автоматическое завершение, предикативный ввод) состоит в том, что почти все они не ориентированы на использование лингвистических ресурсов, прежде всего средств автоматического лингвистического анализа текста при том, что сами алгоритмы индексирования поисковых запросов и текстовых документов в системе информационного поиска основываются именно на таких ресурсах. Не учитываются также типы поисковых запросов и их синтаксических структур, а также само наполнение полнотекстовых баз данных с целью автоматического построения базы подсказок, что существенно сдерживает качественное решение задачи QTA. Реализуемые автором диссертации цели и задачи исследования, включая авторскую разработку метода, алгоритмов и лингвистического обеспечения решения задачи предиктивного автодополнения русскоязычных запросов пользо-

вателей в системах информационного поиска и их реализации в виде промышленного прототипа, выводы и результаты диссертационного исследования, направлены на устранение указанных недостатков, что обеспечит более качественное решение задачи QTA и тем самым облегчит работу пользователей с системами информационного поиска.

Научная новизна диссертационного исследования Ю.Д. Голяк состоит в том, что автору удалось грамотно и аргументировано сформулировать концепцию решения задачи QTA, научная значимость которой определяется ее ориентацией на промышленный характер использования и высокие показатели эффективности. Основу этой авторской концепции составили вопросно-ответная система информационного поиска с естественно-языковым интерфейсом пользователя и русскоязычную поисковую базу данных большого объема, являющуюся поисковым пространством системы информационного поиска, а также использование «истории» не уже осуществленного, а предполагаемого поиска в виде множества подсказок, заранее автоматически распознаваемых в текстовых документах из поисковой базы данных. В отличие от разработанных ранее решений, предложенная автором концепция решения задачи QTA ориентирована не только на последовательное завершение уже набранной части поискового запроса, а еще и на погружение ее в контекст в виде использования соответствующих дополнений в любой части поискового запроса с опорой на глубокий лингвистический анализ текстовых документов.

Бесспорная заслуга автора диссертации состоит в том, что он впервые на основе используемого для исследования текстового материала разработал классификацию основных типов поисковых запросов и их базовых синтаксических структур, определил метод решения задачи QTA, который в отличие от существующих, основан на классификации основных типов поисковых запросов и их базовых синтаксических структур, сводится к автоматическому распознаванию этих структур в поисковой базе данных и выбору их лексических наполнений, которые в совокупности составляют базу подсказок для автодополнения поисковых запросов.

Научную новизну и значимость диссертационного исследования Ю.Д. Голяк определяют также разработанная соискателем структурно-функциональная схема системы QTA автодополнения русскоязычных поисковых запросов в системе информационного запроса и определен ее базовый лингвистический процессор. Причем научная новизна состоит не только в разработке самой схемы, но и в ее функциональности (базового лингвистического анализа поисковой базы данных с целью построения лингвистического индекса текстовых документов в виде баз данных SAO-отношений и автоматического распознавания в ней подсказок с целью формирования баз подставок), ориентированной на сформулированную автором концепцию и предложенный метод решения задачи QTA и основанной на лингвистическом процессоре задачи, включающем БЛП и его расширение.

Научную ценность в диссертации соискателя представляет также разработанные им в виде расширения базового собственное лингвистическое обеспечение и алгоритмы автоматического распознавания в поисковой базе данных подсказок, соответствующих базовым синтаксическим структурам основных типов поисковых запросов с целью их автодополнения. Новизна в данном случае состоит в том, что подобного рода ресурсы впервые автором используются для решения задачи QTA.

Бесспорную научную новизну диссертационного исследования представляет также разработанный автором прототип оригинальной системы QTA/R предиктивного автодополнения русскоязычных поисковых запросов в системе информационного поиска, отличительными чертами которого являются: предоставление пользователю возможности более точно формулировать свою информационную потребность, существенно минимизировать общее время решения задачи, решать задачу QTA, не ориентируясь на «историю» информационного поиска, получать релевантную реакцию поисковой системы. В целях верификации надежности разработанного автором прототипа, соискателем было осуществлено тестирование и внедрение системы в промышленную эксплуатацию, которые показали, что эта система, основанная на предложенных соискателем концептуальных и алгоритмических решений и лингвистических ресурсах, обладает высокими качественными характеристиками.

Автореферат диссертации Ю.Д. Голяк качественно структурирован, в нем нет избыточной информации, все содержательные компоненты логически согласованы и взаимосвязаны, он в точности отражает содержание и композицию диссертационного исследования. К несомненным достоинствам рецензируемого автореферата диссертации можно также отнести умение автора в предельно сжатом виде доносить нужную информацию, что свидетельствует о высокоразвитой способности соискателя к переструктурированию информации с сохранением в ней только наиболее значимых смыслов. Выводы, сформулированные автором диссертации, и результаты, полученные в ходе исследования, обоснованы и убедительны. Обоснованность и достоверность научных положений, выводов и рекомендаций, сформулированных в диссертации, подтверждается достаточным количеством фактов, установленных в ходе исследования, большим объемом исследовательского материала, выбором оптимальных методов исследования, которые соответствуют цели работы, поставленным задачам и материалу исследования. Сформулированные в автореферате диссертации научные положения и выводы основаны на фактических данных, представленных в приведенных в автореферате диссертации таблицах и рисунках.

Серьезных замечаний в автореферате диссертации Ю.Д. Голяк рецензентом не выявлено.

С учетом вышеизложенного считаю, что диссертация Юлии Дмитриевны Голяк «Предиктивное автодополнение запросов пользователей в системах информа-

ционного поиска (на материале русского языка)» соответствует паспорту специальности 10.02.21 – прикладная и математическая лингвистика, отвечает всем требованиям, предъявляемым к кандидатским диссертациям, а ее автор заслуживает присуждения ученой степени кандидата филологических наук по специальности 10.02.21 – прикладная и математическая лингвистика.

Выражаю свое согласие на размещение моего отзыва на автореферат диссертации Голяк Ю.Д. «Предиктивное автодополнение запросов пользователей в системах информационного поиска (на материале русского языка)» на сайте учреждения образования «Минский государственный лингвистический университет».

*Информация о лице, составившем отзыв:*

доктор филологических наук, доцент,  
заведующий кафедрой теории и методики  
преподавания русского языка как иностранного  
филологического факультета

Белорусского государственного университета

(220004, Республика Беларусь, г. Минск,

пр. Независимости, 4,

[bsu@bsu.by](mailto:bsu@bsu.by); сайт: [bsu.by](http://bsu.by))

Лебединский Сергей Иванович

02.03.2023

