

УДК 007:681.512.2

СЕРВИСЫ ПОЛНОТЕКСТОВОГО ПОИСКА В РАСПРЕДЕЛЕННОЙ ИНФОРМАЦИОННОЙ СРЕДЕ (ПРОЕКТ HUMANITARIANA)

С.Х. Ляпин^{a,b}, А.В. Куковьякин^b, И.А. Мбого^a, И.И. Толстикова^a, А.В. Чугунов^a

^a Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

^b ООО «Константа», Архангельск, 163045, Российская Федерация

Адрес для переписки: lyapins@yandex.ru

Информация о статье

Поступила в редакцию 15.10.14, принята к печати 20.12.14

doi: 10.17586/2226-1494-2015-15-1-155-162

Язык статьи – русский

Ссылка для цитирования: Ляпин С.Х., Куковьякин А.В., Мбого И.А., Толстикова И.И., Чугунов А.В. Сервисы полнотекстового поиска в распределенной информационной среде (проект Humanitariana) // Научно-технический вестник информационных технологий, механики и оптики. 2015. Том 15. № 1. С. 155–162

Аннотация

Постановка проблемы. Обосновывается возможность использования сервисов полнотекстового поиска как в универсальных, так и специализированных (по ресурсной базе) электронных библиотеках для извлечения и анализа контекстного знания в гуманитарных дисциплинах. В этой связи описываются архитектура и сервисы виртуального информационно-ресурсного центра для извлечения знаний из гуманитарных текстов, создаваемого в рамках проекта «Humanitariana». Рассматривается функциональная интеграция ресурсов и сервисов полнотекстового поиска в распределенной децентрализованной среде, организованной в архитектуре Интернет/Интранет под управлением клиентского (пользовательского) браузера, обращающегося к множеству независимых серверов. Описан алгоритм осуществления распределенного полнотекстового запроса.

Методы. Используется методика сочетания частотно-ранжированных и абзацно-ориентированных полнотекстовых запросов: первые используются для предварительного анализа предметной области произведения или их совокупности (экспликация «вертикального» контекста, или макроконтекста), вторые – для экспликации «горизонтального» контекста, или микроконтекста в рамках авторского абзаца. При этом результаты частотно-ранжированных запросов используются для составления абзацно-ориентированных запросов.

Результаты. Приведены результаты текстологических исследований по темам: «проблема факта в русской философии», «проблема одиночества в русской философии и культуре». По совокупной ресурсной базе объемом около 2500 полнотекстовых ресурсов по каждой из тем эксплицировано и кратко описано около 50 фрагментов контекстного знания с целью дальнейшего его экспертного изучения.

Практическая значимость. Предлагаемая технология (сервисы продвинутого полнотекстового поиска в распределенной информационной среде) может быть использована для информационной поддержки гуманитарных исследований и гуманитарного образования, для функциональной интеграции ресурсов и сервисов различных организаций, для проведения междисциплинарных исследований.

Ключевые слова: полнотекстовый поиск, экспликация контекстного знания, «горизонтальный» контекст, «вертикальный» контекст, функциональная интеграция ресурсов, децентрализованная распределенная среда, metasearch engine.

Благодарности

Работа выполнена при поддержке гранта Российского гуманитарного научного фонда (грант РГНФ № 14-03-12017). Выражаем благодарность Университету ИТМО (Санкт-Петербург) за возможность использования своей информационно-телекоммуникационной инфраструктуры, а также ООО «Константа» (Архангельск) за предоставление специализированного прикладного программного обеспечения – многофункциональной информационной системы T-Libra с сервисами продвинутого полнотекстового поиска.

SERVICES OF FULL-TEXT SEARCHING IN A DISTRIBUTED INFORMATION ENVIRONMENT (PROJECT HUMANITARIANA)

S. Kh. Lyapin^{a,b}, A.V. Kukovyakin^b, I.A. Mbogo^a, I.I. Tolstikova^a, A.V. Chugunov^a

^a ITMO University, Saint Petersburg, 197101, Russian Federation

^b "Constanta" Ltd., Arkhangelsk, 163045, Russian Federation

Corresponding author: lyapins@yandex.ru

Article info

Received 15.10.14, accepted 20.12.14

doi: 10.17586/2226-1494-2015-15-1-155-162

Article in Russian

Reference for citation: Lyapin S.Kh., Kukovyakin A.V., Mbogo I.A., Tolstikova I.I., Chugunov A.V. Services of full-text searching in a distributed information environment (project Humanitarian). *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2015, vol. 15, no. 1, pp. 155–162 (in Russian)

Abstract.

Problem statement. We justify the possibility of full-text search services application in both universal and specialized (in terms of resource base) digital libraries for the extraction and analysis of the context knowledge in the humanities. The architecture and services of virtual information and resource center for extracting knowledge from the humanitarian texts generated by «Humanitarian» project are described. The functional integration of the resources and services for a full-text search in a distributed decentralized environment, organized in the Internet / Intranet architecture under the control of the client (user) browser accessing a variety of independent servers. An algorithm for a distributed full-text query implementation is described.

Methods. Method of combining frequency-ranked and paragraph-oriented full-text queries is used: the first are used for the preliminary analysis of the subject area or a combination product (explication of "vertical" context, or macro context), the second - for the explication of "horizontal" context, or micro context within copyright paragraph. The results of the frequency-ranked queries are used to compile paragraph-oriented queries.

Results. The results of textual research are shown on the topics "The question of fact in Russian philosophy", "The question of loneliness in Russian philosophy and culture". About 50 pieces of context knowledge on the total resource base of about 2,500 full-text resources have been explicated and briefly described to their further expert investigating.

Practical significance. The proposed technology (advanced full-text searching services in a distributed information environment) can be used for the information support of humanitarian studies and education in the humanities, for functional integration of resources and services of various organizations, for carrying out interdisciplinary research.

Keywords: full-text searching, contextual knowledge explication, "horizontal" context, "vertical" context, functional integration of resources, decentralized distributed environment, metasearch engine.

Acknowledgements. This work was supported by the Russian Humanitarian Science Foundation grant (RHF grant #14-03-12017). We express our gratitude to ITMO University (Saint Petersburg) for the opportunity to use their information and telecommunication infrastructure, as well as to "Constanta", Ltd. (Arkhangelsk) for the provision of application-specific software - Multi-functional Information System T-Libra with advanced full-text search services.

Введение. Полнотекстовый поиск и распределенная среда

Поисковые технологии применительно к текстовым ресурсам используются достаточно давно и в широком диапазоне. При этом до сих пор весьма распространены подходы, при которых прямой и непосредственный пользовательский поиск по полнотекстовым массивам универсальных электронных библиотек является скорее исключением, чем правилом.

Так, библиотекарями активно разрабатывается инструментарий «внешней» (по отношению к содержанию самого полнотекстового ресурса) поддержки осуществления полнотекстового поиска для повышения его эффективности (расширение библиографического описания ресурса, транзакционные ссылки, инструкции по поиску в базах данных и т.п.) [1]. Или, с другой стороны, создаются специализированные программные средства для изучения конкретных полнотекстовых ресурсов – например, философского наследия Людвиг Витгенштейна в формате XML-базы данных. При этом используются основанные на правилах (rule-based) поисковые технологии, совместно с локальной грамматикой, в комбинации с электронным словарем и различными компьютерными средствами для включения лемматизации, поддержки семантических и синтаксических запросов к тексту [2].

Вместе с тем важной тенденцией развития современной информационной среды является непосредственное использование сервисов полнотекстового поиска не только в специализированных информационных системах, но и в универсальных электронных библиотеках: переход от поиска по каталогу (с последующей презентацией найденных ресурсов в виде файлов) к предоставлению пользователю продвинутых поисковых сервисов по полнотекстовым и комбинированным ресурсам. Эта тенденция отмечается исследователями и разработчиками для очень широкого диапазона решаемых задач: от разработки онтологий информационных систем «электронного правительства» [3, 4] и создания коллекций электронных документов по этой тематике [5] до использования сервисов полнотекстового поиска для увеличения цитируемости работ [6] или для поддержки аналитической, образовательной и исследовательской деятельности [7–9].

Другой тенденцией, дополняющей и усиливающей первую, является создание распределенных информационных систем – в весьма широком диапазоне понимания «распределенности», – обладающих в том числе функционалом сквозного, транспарентного для пользователя, полнотекстового поиска [10].

Распределенная среда, создаваемая в той или иной архитектуре, позволяеткратно увеличить количество доступных пользователю ресурсов.

Целью работы является обоснование возможности использования сервисов полнотекстового поиска как в универсальных, так и в специализированных (по ресурсной базе) электронных библиотеках для извлечения и анализа контекстного знания в гуманитарных дисциплинах для информационной поддержки гуманитарных исследований и гуманитарного образования, а также функциональной интеграции ресурсов и сервисов различных организаций для проведения междисциплинарных исследований.

Проект «Humanitarian»

В рамках проекта, рассчитанного на 2014–2016 г.г., создается виртуальный информационно-ресурсный центр для извлечения знаний из гуманитарных текстов и интеграции ресурсов и сервисов в распределенной среде, функционально объединяющей несколько удаленных серверов. Организация «взаимодействия» между этими серверами осуществляется по методологии и технологии Metasearch engine [11], где в качестве машины метапоиска используется функционал пользовательского браузера, обращаясь к множеству независимых и напрямую не связанных между собою серверов.

Для целей проекта производится адаптация функций универсальной электронной библиотеки на основе информационной системы (ИС) T-Libra [12], обладающей возможностями продвинутого полнотекстового поиска и различными формами презентации полученных результатов. Такие технологии позволяют эксплицировать как «горизонтальные» микроконтексты (в пределах авторского абзаца), так и «вертикальные» макроконтексты (в пределах целого документа или совокупности документов), производить тем самым содержательное структурирование предметной области произвольной совокупности документов.

В проекте реализуется также возможность распределенного полнотекстового поиска, функционально объединяющего ресурсы нескольких аналогичных (созданных на базе ИС T-Libra) электронных библиотек. В рамках проекта разрабатывается типология задач автоматизированного извлечения контекстного знания из гуманитарных текстов, создаются методики составления запросов разного типа и вида для типовых задач извлечения знаний, некоторые из них будут реализованы в технологиях запроса в ходе реализации проекта.

На первом этапе (2014 г.) создается экспериментальная информационная среда с вышеназванными возможностями и совокупной ресурсной базой в 1000 русскоязычных гуманитарных текстов (предметные области: философия, история, культурология). Разрабатывается и описывается методология осуществления запросов, ориентированных на экспликацию различных смысловых контекстов и на тематизацию материала. На следующих этапах проекта (2015–2016 г.г.) будет обеспечено дальнейшее развитие интеллектуальных возможностей поиска:

- расширение возможностей полнотекстового поиска в распределенной среде (реализованы частотно-ранжированные и гибридные запросы);
- масштабирование проекта: расширение числа организаций-участников центра, обладающих аналогичными технологиями (до 7–8 организаций сферы образования, науки и культуры) и увеличение доступной совокупной ресурсной базы (до 10 тысяч полнотекстовых ресурсов), по которой будет реализован распределенный полнотекстовый поиск с открытым доступом из Интернет.

В результате выполнения проекта будет впервые создана новая высокотехнологичная научная инфраструктура с открытым доступом из сети Интернет, сервисы которой обеспечат извлечение контекстных знаний из текстов гуманитарной направленности (философия, культурология, история). Функционирование ресурсного центра может осуществляться как в режиме локальной сети в каждой из организаций-участников (технология клиент-сервер), так и в распределенной информационной среде (децентрализованная архитектура в среде Интернет, пользовательский запрос ко всем ресурсам с любого из серверов, либо с сервера-координатора, выбранного участниками проекта).

Архитектура информационной системы и базовые сервисы полнотекстового поиска

Архитектура. Используемая в проекте ИС T-Libra функционирует в клиент-серверной Интернет/Интранет архитектуре. На стороне пользователя предполагается лишь наличие Интернет-браузера и стандартных прикладных программ по работе с файловыми ресурсами. На стороне сервера – операционная система Windows, СУБД MySQL, Веб-сервер Apache, сервер приложения. Вся бизнес-логика вынесена в сервер приложения. В настоящее время на серверах участников проекта используются Linux-сервера с развернутой на них виртуальной машиной Windows. В дальнейшем планируется переход на чистую Linux-версию ИС T-Libra.

Сервисы полнотекстового поиска. В используемой нами версии электронной библиотеки имеются следующие типы полнотекстового поиска:

- абзацно-ориентированный;
- частотно-ориентированный.

При этом абзацно-ориентированный поиск представлен разновидностями работы как в локальной, так и в распределенной среде.

Абзацно-ориентированный поиск предназначен для поиска и презентации текста с точностью до отдельных авторских абзацев, содержащих заданную пользователем терминологическую структуру (тем самым эксплицируется «горизонтальный» микроконтекст, в котором в составе абзаца находятся искомые термины). Авторский абзац выбран в качестве естественной единицы смыслового членения текста. Обеспечивается поддержка нескольких видов и различных форм презентации результатов этого поиска.

- **Простой («однослойный») тематический поиск**, с одним комплексным полем для ввода терминов и использованием для этих терминов операторов логического объединения, обязательного исключения или обязательного включения термина в запрос. Результатом поиска является список абзацев, удовлетворяющих заданным условиям.

Каждый из абзацев, входящих в результаты запроса, может быть одним «кликом» мышки раскрыт до своего полного вида. Используя опцию «Контекст» в левом меню, можно последовательно раскрыть абзацы до и после найденного – вплоть до кластера из семи абзацев (три абзаца «до», три абзаца «после», плюс сам абзац – результат запроса).

Имеется возможность посмотреть с этой же экранной страницы соответствующий ресурс (статью, книгу и т.д.) в файловом виде; ресурс при этом может быть представлен в различных формах – текстового документа, графического образа документа (важно для архивных ресурсов), сопровождающего документ аудио- или видеофайла (важно для организации электронных выставок и коллекций).

Имеется также возможность оценки пользователем найденных абзацев и автоматизированной сборки выбранных абзацев с помощью опции «Собрать тему». В результате пользователь получает файл, в котором собраны тематически ориентированные абзацы из различных документов электронной библиотеки и указаны ресурсы (вместе с их библиографическими описаниями), откуда взяты эти абзацы.

- **Расширенный («многослойный») тематический поиск**. Этот вид поиска содержит функционал дополнительной тематической фокусировки запроса. Соответствующий инструментарий включает в себя:

- формирование нескольких поисковых полей («слоев»);
- включение в запрос дополнительных количественных параметров его фокусировки.

Поисковое поле «слои» представляет собой технический инструмент для выделения того или иного содержательного «аспекта» интересующей пользователя «темы»; всего может быть сформировано от 2 до 8 слоев. Например, в первом слое вводим термин «факт», во втором – термин «разум», в третьем – термин «истина». Тем самым в структуре запроса тематика «факта» специализирована (аспектуализирована) в связи с «истиной» и «разумом».

Между слоями действует операция логического пересечения; внутри слоя – операция логического объединения заданных терминов. Имеется возможность комбинировать актуально используемые слои, например, из трех слоев сделать какие-либо два обязательными, а один – произвольным выбираемым при осуществлении запроса.

Еще более точная тематическая фокусировка запроса достигается за счет выполнения дополнительных условий:

- указания минимально необходимого количества поисковых слоев (от 2 до 8);
- указания максимального расстояния между терминами, принадлежащими разным слоям: от 0, когда слова из двух разных слоев запроса в составе абзаца примыкают друг к другу (например, «факт истории» и т.д.), до произвольной величины.

В рамках распределенной информационной среды в настоящее время реализованы абзацно-ориентированные запросы, наиболее востребованные для тематической обработки документов.

Примеры экспликации контекстного знания с помощью сервисов полнотекстового поиска.

Развивая идею о сервисах полнотекстового поиска для поддержки гуманитарных исследований, на протяжении нескольких последних лет некоторыми из авторов настоящей работы в рамках различных мероприятий были осуществлены предварительные тематические исследования, в частности, по таким крупным темам, как «Проблема факта в русской философии», «Проблема одиночества в русской философии и литературе», «Проблема человека в русской философии и культуре». Эти и аналогичные исследования будут существенно расширены – как по привлекаемой ресурсной базе, так и по изучаемым темам, их аспектам и используемым сервисам – в рамках проекта «Humanitarianiana».

Для экспликации вышеназванных концептов применялась методика сочетания частотно-ориентированных запросов с абзацно-ориентированными. Для этого вначале строились соответствующие частотно-ранжированные списки терминов (существительных) – «терминограммы» – относительно исследуемых терминов (соответственно: «факт», одиночество», «человек»). При этом на заданную глубину (как правило, в 10 слов) средствами относительного частотного поиска автоматически создавались таблицы частотной встречаемости терминов по выбранной совокупности ресурсов. В итоге мы получали экспликацию предметной области, коррелирующей с исследуемыми терминами/концептами (соответственно: «факт», одиночество», «человек») – т.е. экспликацию «вертикального контекста». Затем термины из этих таблиц включались в абзацно-ориентированные запросы, что позволяло эксплицитировать микроконтекст («горизонтальный контекст») употребления этих терминов в рамках авторского абзаца.

Экспликация концепта фактуальности в философии Н.А. Бердяева. С помощью относительного частотного поиска (относительно термина «факт») можно предварительно выявить предметную

область (смысловой макроконтекст), в котором существует концепт «факт» в философских текстах Н.А. Бердяева в целом или в тексте отдельного произведения [13]. В частности, для исследуемых работ Бердяева, содержащих термин «факт» (24 произведения), в смысловую окрестность «факта» попадают (т.е. встречаются в абзацах, где обязательно имеется и термин «факт»), в порядке убывания частотности, следующие термины: 1. «жизнь» (850 раз), 2. «мир» (811 раз), 3. «человек» (788 раз), 4. «дух» (484 раза), 5. «Бог» (483 раза), 6. «сознание» (401 раз), 7. «свобода» (377 раз), 8. «факт» (366 раз), 9. «философия» (324 раза), 10. «история» (311 раз).

Опираясь на эти результаты, эксплицирующие предметную область фактуальности в работах Н.А. Бердяева, можно далее с помощью абзацно-ориентированного поиска осуществить экспликацию горизонтальных микроконтекстов (в пределах авторского абзаца), коррелирующих с концептом «факт».

Можно использовать многослойный (например, трехслойный) запрос: в один слой ввести термин «факт», в два других – термины из бердяевской терминограммы (например, «мир» и «человек»), выбрать дополнительные параметры фокусировки запроса (все 3 слоя актуальны, расстояние между терминами не более 8 слов) и получить результат – 7 абзацев в 4 произведениях (поиск произведен по массиву из 64 текстов Бердяева) (рис. 1).

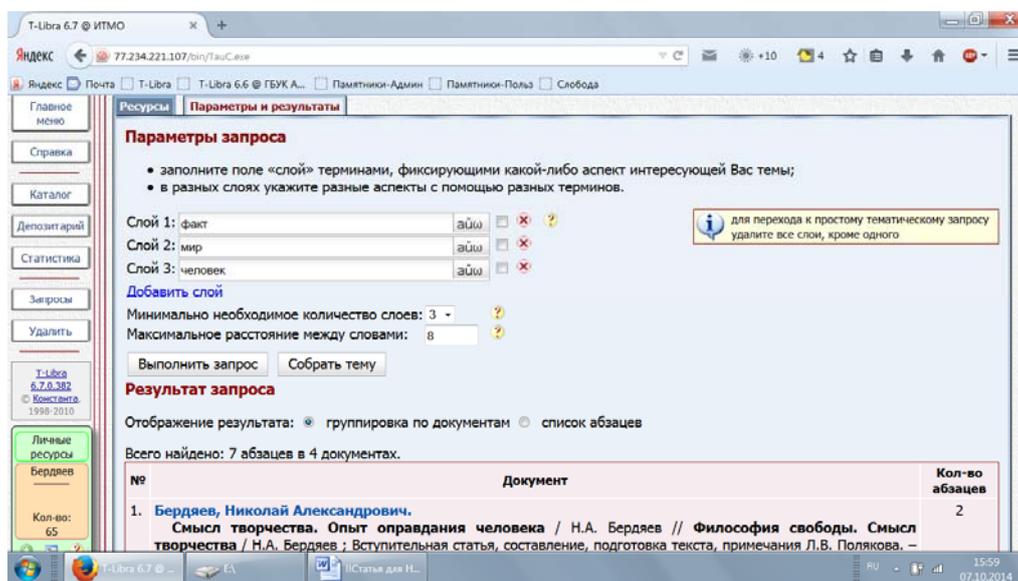


Рис. 1. Многослойный (трехслойный) запрос с терминами [«факт», «мир», «человек»] по массиву текстов Н.А. Бердяева

Экспертная оценка этого и других найденных абзацев говорит о высокой степени релевантности результатов таким образом организованного полнотекстового запроса. Эксплицируются микроконтексты концепта фактуальности, связанные с такими аспектами темы: «гуманизм и факт рабства человека у природного мира»; «факт самосознания человека и трансцендентность сознания»; и т.д.

Выбирая другие термины из терминограмм Бердяева (например, «факт» + «сознание» + «свобода»; или «факт» + «философия» + «история» и т.д.), включая их в абзацно-ориентированные запросы и варьируя условия их фокусировки, можно эксплицировать самые разные семантические микроконтексты в рамках разработки темы «Проблема факта в философии Бердяева».

С соответствующими изменениями сказанное относится к другим авторам – например, терминограммам Лосева и Канта и соответствующим абзацно-ориентированным запросам по их произведениям. Все это дает, в свою очередь, текстологический материал для гносеологической компаративистики.

Полнотекстовый поиск в децентрализованной распределенной среде, управляемой пользовательским браузером

При выборе модели организации распределенной среды для сервисов полнотекстового поиска мы ориентировались на Веб-сервисы и Интернет-протоколы. Это связано как с архитектурой ИС T-Libra (клиент-серверная архитектура, в которой в качестве клиента выступает Веб-браузер), так и с очевидной необходимостью учитывать тенденции развития современного информационного пространства.

В этой связи была выбрана модель децентрализованной среды под управлением пользовательского браузера, обращающегося к множеству независимых серверов, находящихся в ведении различных организаций. Этот подход вполне укладывается в парадигму распределенных информационных систем [14], достаточен для наших целей и позволяет не рассматривать более сложные варианты построения

распределенных систем, связанные, например, с взаимодействием унифицированного протокола Z39.50 [15] с Интернет-протоколами [16] (рис. 2).

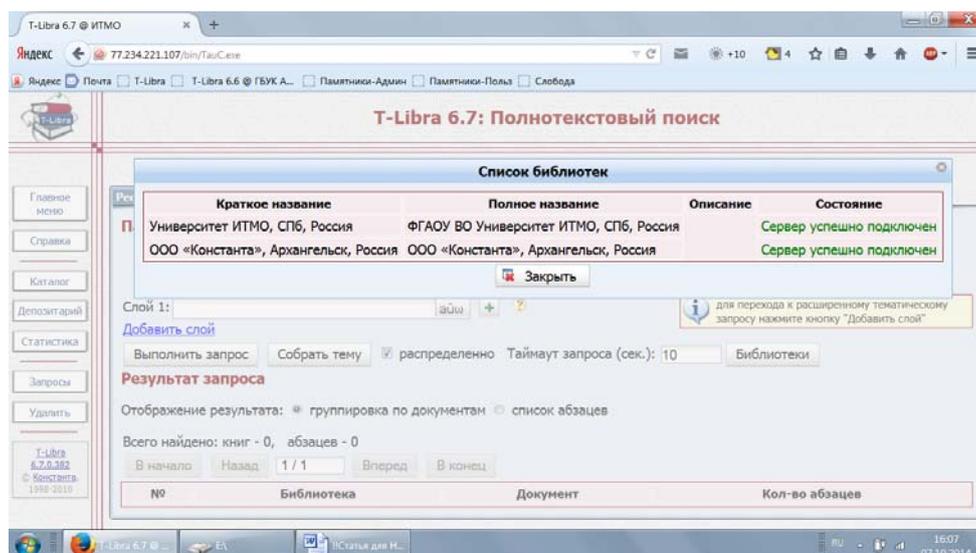


Рис. 2. Сформирован распределенный полнотекстовый запрос

С базового сервера (Университет ИТМО) сформирован распределенный полнотекстовый запрос, на который откликнулся также сервер другого участника проекта – ООО «Константа». Это показано на соответствующей странице во всплывающем окне.

Наш подход может быть рассмотрен как вариант Metasearch engine. Инструментом машины мета-поиска является в этом случае пользовательский браузер. Прямого взаимодействия серверов при этом не требуется.

Алгоритм организации распределенного запроса

Замечание. «Базовый» сервер – это сервер, который сформировал страницу, на которой пользователь нажал кнопку «Создать распределенный запрос». Базовым он является функционально, т.е. для этого конкретного запроса. Практически это тот сервер, с которым пользователь начал работу в одной из библиотек, включенных в распределенную среду.

1. клиентская часть ИС T-Libra (функционирующая в браузере пользователя), работающая с каким-либо сервером T-Libra («базовый» сервер), получает от базового сервера список адресов других серверов, которые будут участвовать в полнотекстовом запросе (этот список ведется администратором базового сервера T-Libra).
2. после заполнения формы полнотекстового запроса пользователь инициирует выполнение запроса (кнопкой «Выполнить запрос»).
3. браузер рассылает http-запросы (на выполнение полнотекстового запроса) на базовый сервер и прочие, перечисленные в списке.
4. каждый сервер выполняет запрос, формирует результат и отправляет в клиентскую часть 10 лучших ответов и общий размер результата.
5. из полученных частичных «топ 10» (от каждого сервера) клиентская часть гарантированно строит и отображает пользователю «топ 10» ответа в целом (первая страница результата) и суммарный размер результата.
6. в ожидании реакции пользователя клиентская часть рассылает http-запросы (на пересылку остатков результата) на все серверы, участвующие в выполнении запроса.
7. клиентская часть принимает ответы от них, объединяя частичные результаты и позволяя пользователю просмотреть следующих страниц результата распределенного полнотекстового запроса.

Заключение. Развитие проекта «Humanitarianiana»

Развитие ресурсно-сервисной среды, создаваемой в рамках проекта «Humanitarianiana», предполагается в нескольких взаимно дополняющих друг друга направлениях:

- обеспечение многоплатформенности (адаптация для операционной системы семейства Linux в серверной части информационной системы);
- развитие интеллектуальности поиска (многопараметрическая кластеризация результатов запроса, разработка и использование тезаурусов в полнотекстовом поиске, разработка и реализация каскадных и

- гибридных запросов, семантический анализ и концептуальное моделирование текста, внедрение технологий программных агентов и мультиагентных систем);
- масштабирование распределенной среды: функциональная интеграция со специализированными информационными системами (с полнотекстовой электронной библиотекой, с распределенными информационными системами, обладающими соответствующим функционалом, с веб-сайтами с материалами научных конференций).

References

1. Fagan J.C. The effects of reference, instruction, database searches, and ongoing expenditures on full-text article requests: an exploratory analysis. *Journal of Academic Librarianship*, 2014, vol. 40, no. 3–4, pp. 264–274. doi: 10.1016/j.acalib.2014.04.002
2. Hadersbeck M., Pichler A., Fink F., Gjesdal Ø.L. Wittgenstein's nachlass: WiTTFind and wittgenstein advanced search tools (WAST). *Proc. 1st Int. Conf. on Digital Access to Textual Cultural Heritage, DATeCH 2014*. Madrid, Spain, 2014, pp. 91–96. doi: 10.1145/2595188.2595202
3. Yudina T.N., Bogomolova A.V. UIS ROSSIYA: ontologiya predmetnoi oblasti "gosudarstvennoe upravlenie". *Trudy XIV Vserossiiskoi Konferentsii "Internet i Sovremennoe Obshchestvo", IMS-2011* [Proc. XIV All-Russian Conf. "Internet and Modern Society", IMS-2011]. St. Petersburg, Russia, 2011, pp. 225–230.
4. Heeks R., Santos R. *Understanding Adoption of e-Government: Principals, Agents and Institutional Dualism*. Available at: www.sed.manchester.ac.uk/idpm/research/publications/wp/igovernment/igov_wp19.htm (accessed 30.11.2014).
5. Mbogo I.A., Chugunov A.V. Elektronnaya kolleksiya "Elektronnoe gosudarstvo": tekhnologicheskie aspekty [Electronic collection "Electronic government": technological aspects]. *Trudy XV Vserossiiskoi Ob"edinennoi Konferentsii "Internet i Sovremennoe Obshchestvo"* [Proc. XV All-Russian Conference "Internet and Modern Society"]. St. Petersburg, Russia, 2012, pp. 345–347.
6. Lyapin S.Kh., Kukovyakin A.V. Servisy polnotekstovogo poiska kak instrument povysheniya tsitiruемости nauchnykh rabot i reitinga VUZa [Full-text search services as a tool to improve the citation of scientific papers and university rankings]. *Trudy XX Vserossiiskoi Nauchno-Metodicheskoi Konferentsii Telematika'2013* [Proc. XX Scientific and Technical Conference Telematika'2013]. St. Petersburg, Russia, 2013, vol. 1, pp. 15–17.
7. Lyapin S.Kh. Elektronnaya polnotekstovaya biblioteka dlya podderzhki sotsiogumanitarnykh issledovaniy [Electronic full-text library to support social and humanitarian studies]. *Trudy XX Vserossiiskoi Nauchno-Metodicheskoi Konferentsii Telematika'2013* [Proc. XX Scientific and Technical Conference Telematika'2013]. St. Petersburg, Russia, 2013, vol. 2, pp. 317–318.
8. Lyapin S.Kh. Servisy elektronnoi polnotekstovoi biblioteki dlya obrazovaniya, nauki i kul'tury [Electronic full-text library services for education, science and culture]. *Nauchnaya Periodika: Problemy i Resheniya*, 2013, no. 2(14), pp. 9–17.
9. Lyapin S.Kh., Kukovyakin A.V. Elektronnaya polnotekstovaya biblioteka dlya podderzhki analiticheskoi i issledovatel'skoi deyatel'nosti [Electronic full-text library to support analytical and research activities]. *Trudy XVI Vserossiiskoi ob"edinennoi konferentsii "Internet i sovremennoe obshchestvo", IMS-2013* [Proc. XVI All-Russian Conf. "Internet and Modern Society", IMS-2013]. St. Petersburg, Russia, 2013, pp. 163–170.
10. Lyapin S.Kh. Kak proiti v raspredelennuyu biblioteku? [How to pass in the distributed library?]. *Sovremennaya nauka: aktual'nye problemy teorii i praktiki. Seriya: Gumanitarnye nauki*, 2012, no. 7–8, pp. 17–21.
11. *Metasearch Engine*. Available at: http://en.wikipedia.org/wiki/Metasearch_engine (accessed 30.11.2014).
12. *T-Libra 6.7*. Available at: <http://demo.tlibra.ru> (accessed 30.11.2014).
13. Lyapin S.Kh. *Teksty, konteksty, kontsepty: ispol'zovanie sovremennykh sistem polnotekstovogo poiska dlya filosofskikh issledovaniy (na materiale russkoi filosofii)* [Texts, contexts, concepts: the use of modern systems of full-text search for philosophical studies (based on Russian philosophy)]. Available at: <http://www.losev-library.ru/index.php?pid=6872> (accessed 30.11.2014).
14. Tanenbaum A.S., van Steen M. *Distributed Systems. Principles and Paradigms*. Prentice Hall, 2002.
15. Zhizhimov O.L., Mazov N.A. Model' raspredelennoi informatsionnoi sistemy Sibirskogo otdeleniya RAN na baze protokola Z39.50. *Elektronnye Biblioteki*, 1999, vol. 2, no 2, p. 12.
16. Maksimov N.V., Sysoikina M.A. O realizatsii elektronnoi biblioteki s ispol'zovaniem protokolov HTTP i Z39.50 [On the implementation of digital library using HTTP and Z39.50]. *Elektronnye Biblioteki*, 2002, vol. 5, no. 1, p. 4.

Ляпин Сергей Хамзеевич

– кандидат философских наук, доцент, научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация; генеральный директор, ООО «Константа», Архангельск, 163045, Российская Федерация, lyapins@yandex.ru

- Куковьякин Алексей Валентинович** – заместитель генерального директора, ООО «Константа», Архангельск, 163045, Российская Федерация, magicmagus@yandex.ru
- Мбого Ирина Анатольевна** – программист, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, irina.mbogo@gmail.com
- Толстикова Ирина Ивановна** – кандидат философских наук, доцент, зав. кафедрой, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, tolstikova_irina@mail.ru
- Чугунов Андрей Владимирович** – кандидат политических наук, доцент, директор Центра, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, chugunov@egov-center.ru
- Sergei Kh. Lyapin** – PhD, Associate professor, scientific researcher, ITMO University, Saint Petersburg, 197101, Russian Federation; CEO, "Constanta" Ltd., Arkhangelsk, 163045, Russian Federation, lyapins@yandex.ru
- Alexei V. Kukovyakin** – Deputy director, "Constanta" Ltd., Arkhangelsk, 163045, Russian Federation, magicmagus@yandex.ru
- Irina A. Mbogo** – programmer, ITMO University, Saint Petersburg, 197101, Russian Federation, irina.mbogo@gmail.com
- Irina I. Tolstikova** – PhD, Associate professor, Department Head, ITMO University, Saint Petersburg, 197101, Russian Federation, tolstikova_irina@mail.ru
- Andrei V. Chugunov** – PhD, Associate professor, Director of the eGovernment Center, ITMO University, Saint Petersburg, 197101, Russian Federation, chugunov@egov-center.ru