

## Технологии извлечения знаний на службе научно-исследовательской деятельности в вузе

Статья рекомендована Б.В. Кристалным 18.11.2016.



**КОНОНОВА Ольга  
Витальевна**

*Кандидат экономических наук, доцент кафедры управления государственными информационными системами Санкт-Петербургского национального исследовательского университета информационных технологий, механики и оптики (Университет ИТМО)*



**КРУТЬКО Елена Алексеевна**

*Ведущий специалист-аналитик отдела развития информационных систем ЖКХ и дежурных служб СПб ИАЦ, магистрант кафедры управления государственными информационными системами Университета ИТМО, Санкт-Петербург*



**ЛЯПИН Сергей Хамзеевич**  
*Научный сотрудник центра технологий электронного правительства факультета технологического менеджмента и инноваций Университета ИТМО, Санкт-Петербург, старший научный сотрудник Архангельского краеведческого музея, Архангельск*

### Аннотация

В статье рассматривается подход к организации научно-исследовательской работы в вузе с использованием информационных технологий извлечения контекстных знаний. Предложена методика использования сервисов контекстного поиска и извлечения контекстных знаний из текстов электронной библиотеки «Humanitarian» — виртуального информационно-ресурсного центра с возможностями функциональной интеграции ресурсов и сервисов в распределенной среде.

Опыт организации исследовательской и аналитической работы на базе сервисов ЭБ «Humanitarian» впервые выносятся на широкое обсуждение. Он систематизирован и представлен в рамках общей концепции развития мегафакультетов Университета ИТМО и дополняет системное понимание организационно-управленческой траектории и технологий информационного общества, способствующих ее воплощению.

### Ключевые слова:

**технологии информационного общества, научно-исследовательская деятельность, технологии извлечения контекстного знания, контекстный поиск, абзацно-ориентированный запрос, частотно-ранжированный запрос.**

Особенностью, характеризующей глобальное информационное общество, является изменение требований к потребительским качествам информации, способам и инструментам ее получения и анализа. Под удовлетворением информационных потребностей понимается прежде всего обеспечение открытости и достоверности данных, а также возможности самостоятельно управлять информационными ресурсами — поиском, анализом, оценкой информации и получением благодаря этому новых экспертных знаний. Именно качество и полнота удовлетворения информационных потребностей социума выступает фактором повышения эффективности управления предприятиями, организациями, государственным сектором экономики и государственными институтами, деятельностью отдельного человека.

Задача повышения эффективности управления может быть решена за счет использования механизмов извлечения контекстного знания и методов контекстного поиска из доступного пользователям

информационного массива [1–3]. Под *контекстным знанием* подразумевается умение правильно воспринимать контекст, извлекать и интерпретировать профессионально значимую информацию из любых источников. *Контекстный поиск* — это метод последовательного поиска фрагментов текстовых записей, релевантных пользовательскому запросу в соответствии с требуемым контекстом.

Методы контекстного поиска включены в большинство современных VI-систем и предназначены для поддержки проведения бизнес-аналитики, обеспечения свободного поиска и интеллектуальной выборки данных в систематизированных хранилищах информации, повышения гибкости бизнеса. Важным аспектом применения методов контекстного поиска является поддержка научно-исследовательской деятельности в организациях и компаниях [4]. На практике использование методов контекстного поиска требует перестройки мировоззрения сотрудников и креативного подхода к изменению критериев сбора, обработки, анализа и оценки прагматической ценности информации. Поэтому для формирования внутренней мотивации, соответствующих компетенций, приобретения устойчивых навыков использования технологий контекстного поиска следует внедрять данные технологии в учебный процесс вузов, в первую очередь по программам магистратуры в рамках научно-исследовательской работы студентов. Для достижения успеха начинания необходимо, чтобы методы и технологии контекстного поиска использовались всем научно-преподавательским составом вуза (на первых порах — отдельного факультета или кафедры) в проектной, аналитической и образовательной деятельности.

Предпринимая исследование об использовании информационных технологий при организации НИР в вузе, мы ставили перед собой следующие задачи:

- проанализировать структуру, этапы и последовательность выполнения научно-исследовательской работы в вузе, выделить потребности каждого из этапов;
- выбрать эффективный инструментарий извлечения и анализа контекстных знаний, обеспечивающий доступность данных и открытость исследований;
- разработать методику организации НИР с использованием данного инструментария.

#### **Организация научно-исследовательской деятельности в вузе.**

Содержание проектной, аналитической, образовательной и научно-исследовательской деятельности в вузе предполагает написание рефератов, обзоров, аналитических отчетов, подготовку научных публикаций, учебных кейсов, списков источников и ключевых слов. Кроме того, НИР в вузе подразумевает такую организацию работы студентов и сотрудников, которая обеспечивала бы обоснованный выбор тематики, а затем конкретных тем исследований с доказательством ее актуальности.

Научно-исследовательская работа магистрантов и аспирантов нацелена на формирование общекультурных и общепрофессиональных компетенций, соответствующих ряду задач профессиональной деятельности:

- расширению теоретического кругозора и научной эрудиции;

- овладению методами научного познания, творческого освоения учебного материала;
- формированию навыков анализа и обобщения результатов НИР с использованием современных достижений науки и техники;
- формированию навыков работы с научной литературой;
- овладению методами анализа и управления информационными ресурсами;
- овладению навыками подготовки публикаций по тематике НИР.

Каждая из перечисленных задач связана с обработкой информационных ресурсов — документов или массивов документов в информационных системах (библиотеках, архивах, фондах, банках данных и т.д.) [5]. Поиск, анализ и отбор полезной информации для научных сотрудников, преподавателей и студентов — актуальная задача, требующая определенных компетенций и навыков. Оптимизация процессов отбора и обработки больших массивов информационных ресурсов необходима и при организации проектной и аналитической работы на кафедре.

При поиске информации накапливается огромное количество исходного материала, который необходимо сохранить и в дальнейшем иметь возможность его обрабатывать и анализировать. Возникает необходимость создавать личные полнотекстовые тематические коллекции научных материалов. Для формирования таких коллекций в настоящее время применяется несколько подходов и инструментов:

- сохранение web-страницы на своем компьютере в исходном формате с использованием возможностей Web-браузера;
- копирование URL найденного текста и дальнейшее сохранение полученного документа в текстовом формате, например в формате Word-документа, а также в некоторой папке на своем компьютере;
- использование систем типа NetSnippets для копирования и сохранения найденных ресурсов с возможностью дальнейшего их поиска.

Однако эти подходы обладают существенными недостатками, среди которых:

- разнородность форматов представляемых данных;
- невозможность создания единообразных библиографических описаний для всех найденных как в открытом интернете, так и в автономных электронных библиотеках разнородных информационных ресурсов;
- невозможность организации эффективного полнотекстового поиска требуемого материала в достаточно больших полнотекстовых коллекциях;

- невозможность представить и индексировать полнотекстовые документы в том виде, который допускает их дальнейшее использование.

Повышение управляемости, качества и интенсификация всех этих видов деятельности могут быть достигнуты за счет использования распределенных информационных систем, обладающих функционалом сквозного, прозрачного для пользователя, полнотекстового контекстного поиска. Такие системы предоставляют следующие возможности:

- содержательное структурирование предметной области произвольной совокупности документов;
- типологизацию задач автоматизированного извлечения контекстного знания из текстов;
- методы контекстного поиска;
- методики составления запросов разного типа и вида для типовых задач извлечения знаний;
- различные формы презентации результатов поиска;
- экспликацию различных смысловых контекстов и тематизацию материала.

Для повышения качества результатов НИР магистрантов и активизации проектной деятельности на факультете ФТМИ ИТМО было принято решение использовать электронную библиотечную информационную систему «Humanitarianana» с функциями гибкого тематизируемого полнотекстового поиска и возможностями извлечения контекстных знаний из текстов при различных формах их презентации (книга, архивный документ и т.д.).

Проект создания электронной библиотеки «Humanitarianana» поддержан грантом РГНФ № 14-03-12017 и рассчитан на 2014–2016 гг. В его рамках при координирующей роли Университета ИТМО (г. Санкт-Петербург) поэтапно создавалась междисциплинарная информационная распределенная среда с открытым доступом, разрабатывалась типология задач автоматизированного извлечения контекстного знания из научных текстов, создавались запросы разного типа и вида для типовых задач извлечения знаний.

Электронная библиотека (ЭБ) «Humanitarianana» разработана в клиент-серверной Интернет/Интранет архитектуре: Web-browser / Web-server + ApplicationServer / Relational DBMS, с протоколами HTTP, CGI, PIPE API, ODBC. Функционирует в среде Windows: на сервере — Windows 2000/2003/XP/Vista/Windows 7, СУБД MySQL, веб-сервер Apache, сервер приложения. На серверах также может функционировать операционная система Linux. Сервер приложения охватывает всю бизнес-логику системы.

На клиентском месте — любая из версий Windows. Клиентской программой для ЭБ «Humanitarianana» является стандартный веб-браузер (поддерживаются MS Internet Explorer, Mozilla Firefox, Opera, Apple Safari, Google Chrome, Яндекс. Браузер). С учетом тенденций развития современного информационного пространства была выбрана модель децентрализованной среды под управлением

пользовательского браузера и с ориентацией на веб-сервисы и интернет-протоколы. Браузер обращается к множеству независимых серверов, находящихся в ведении различных организаций. Прямого взаимодействия серверов при этом не требуется [6].

ЭБ «Humanitarian» может функционировать:

- в режиме локальной сети в каждой из организаций-участников (технология клиент–сервер);
- в режиме распределенной информационной среды (децентрализованная архитектура в среде интернет, пользовательский запрос ко всем ресурсам осуществляется с любого из серверов).

Сервисы ЭБ обеспечивают два вида абзацно-ориентированного и четыре вида частотно-ориентированного контекстного поиска, а также поддерживают различные формы представления результатов запроса.

*Абзацно-ориентированный поиск* — это поиск, осуществляемый по выбранной совокупности ресурсов с учетом словоизменения поисковых терминов (например, множественное или единственное число в английском языке, склонение по падежам в русском языке). Эта методика предназначена 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.

Информационной системой обеспечивается два вида абзацно-ориентированного контекстного запроса (АОЗ) и поддерживаются различные формы представления результатов поиска. Простой («однослойный») контекстный запрос — поиск с одним комплексным полем для ввода терминов и использованием для этих терминов операторов логического объединения, обязательно исключения или обязательного включения термина в запрос. Результатом поиска является список абзацев, удовлетворяющих заданным условиям. Доступные сервисы:

- просмотр с той же экранной страницы соответствующего ресурса (статьи, книги и т.д.) в файловом виде; форматы файлов — текстовые документы, графический образ документа (для архивных ресурсов), аудио- или видеодокументы (для организации электронных выставок и коллекций);
- экспертная оценка пользователем найденных абзацев;
- автоматизированная сборка в отдельный файл тематически ориентированных абзацев из различных документов электронной библиотеки вместе с их библиографическими описаниями и информацией об источниках;
- запись файла на переносной носитель информации.

Расширенный («многослойный») контекстный запрос — поиск с функционалом дополнительной тематической фокусировки запроса. Поисковое поле «слой» (от 2 до 8 слоев) представляет собой технический инструмент для выделения того или иного содержательного «аспекта» интересующей пользователя «темы». Например, в первом слое вводим термин «правительство», во втором — «услуги», в третьем — «регион». Тем самым в структуре запроса тематика «правления» специализирована в связи с «услугами» и «регионом». Еще более точная тематическая фокусировка запроса достигается за счет 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, «электронное правительство» или «государственные услуги» и т.п.), до произвольной величины.

The screenshot shows a search interface with the following sections:

- Параметры запроса:**
  - Instructions: fill the 'layer' field with terms, or use different terms for different layers.
  - Layer 1: Человек
  - Layer 2: Развитие
  - Layer 3: время
  - Layer 4: Проблема
  - Buttons: 'Добавить слой', 'Выполнить запрос', 'Собрать тему', 'распределенно', 'Таймаут запроса (сек.): 10', 'Библиотеки'
- Результат запроса:**
  - Display: 'группировка по документам' selected.
  - Found: 47 books, 87 paragraphs.
  - Navigation: 'В начало', 'Назад', '1 / 5', 'Вперед', 'В конец'
- Table of results:**

№	Библиотека	Документ	Кол-во абзацев
1	Университет ИТМО, СПб, Россия	<b>ОБРАЗ КИБЕРЧЕЛОВЕКА В СОВРЕМЕННОЙ НАУКЕ И КУЛЬТУРЕ</b> / Баева Л. В. – 2015. – С. 56-69. Термины индексирования   Аннотация   Ресурсы	11
2	Университет ИТМО, СПб, Россия	<b>Бакалавр.</b> – 2015. – 73 с. Ресурсы	8
3	Университет ИТМО, СПб, Россия	<b>Изучение проблемы зависимости от компьютерных игр на основе логико-лингвистической модели</b> / В. А. Спесивцев, С. А. Спесивцев. – 2014. – с.274-278. Ресурсы	5

Рис. 1. Многослойный АОЗ с терминами «человек», «развитие», «время», «проблема» по массиву текстов «Компьютерные игры» за 2014–2015 гг. (всего 264 статьи). При заданных условиях фокусировки запроса найдено 87 релевантных абзацев в 47 статьях. Запрос произведен по ЭБ «Humanitarian» Университета ИТМО 23.03.2016

*Частотно-ориентированный поиск* предназначен для построения частотно-ранжированных списков терминов (существительных) и тем самым экспликации различных «вертикальных» макроконтестов, неявно присутствующих в отдельном документе или массиве документов [6]. ЭБ обеспечивает поддержку двух видов частотно-ориентированного запроса (ЧОЗ):

- абсолютный частотный, результатом которого является частотно-ранжированный список существительных, входящих в ресурсы области поиска и приведенных к нормальной форме (именительный падеж, единственное число);
- относительный частотный, результатом которого является частотно-ранжированный список существительных, входящих только в те абзацы, которые содержат заранее заданный пользователем термин.

Результатом частотно-ориентированных запросов являются «терминограммы». Терминограмма — это совокупность единиц анализа (таблица

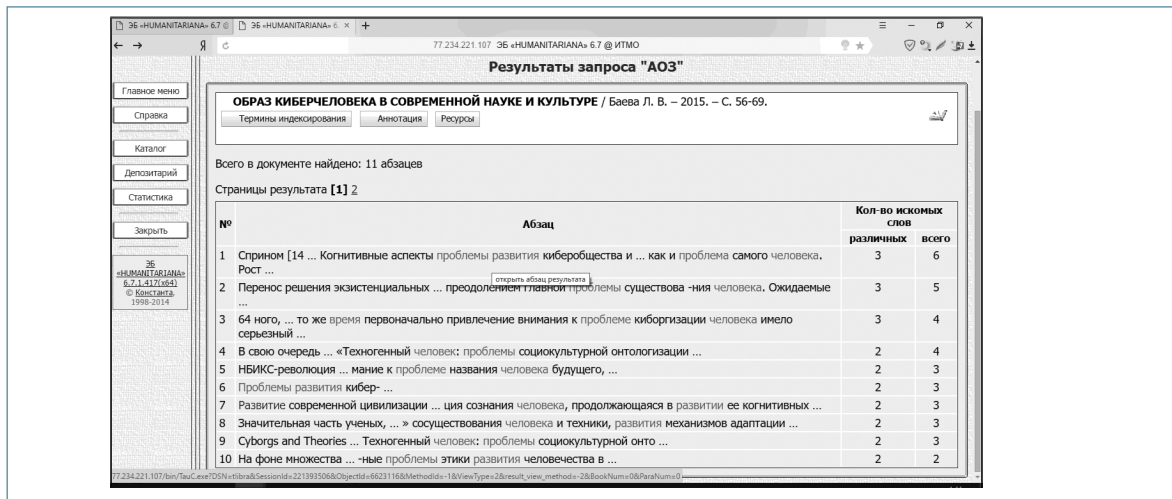


Рис. 2. Результаты многослойного АОЗ с терминами «люди», «развитие», «время», «проблема» по массиву текстов «Компьютерные игры» за 2014–2015 гг. (найденно 36 статей) с раскрытием одной из статей до абзацев. Запрос произведен по ЭБ «Humanitarian» Университета ИТМО 23.03.2016

терминов), содержащая информацию об абсолютной (в числах) и относительной (в промилле,‰) частоте употребления терминов в тексте. Пример результата запроса абсолютного частотно-ориентированного поиска приведен на рисунке 3.

Абсолютный частотный запрос				Абсолютный частотный запрос				Абсолютный частотный запрос			
2005-2007 г.				2008-2012 г.				2013-2015 г.			
27 документов				340 документов				334 документа			
№	Слово	Кол-во	Частота (%)	№	Слово	Кол-во	Частота (%)	№	Слово	Кол-во	Частота (%)
1	процесс	545	2.212	1	игра	5334	8.575	1	игра	7949	8.856
2	система	516	2.094	2	процесс	889	1.429	2	развитие	1555	1.732
3	развитие	479	1.944	3	мир	791	1.272	3	дети	1462	1.629
4	группа	447	1.814	4	дети	791	1.272	4	процесс	1336	1.488
5	игра	404	1.640	5	развитие	773	1.243	5	обучение	1290	1.437
6	работа	378	1.534	6	деятельность	742	1.193	6	работа	1250	1.393
7	уровень	370	1.502	7	время	707	1.137	7	мир	1233	1.374
8	студент	354	1.437	8	человек	678	1.090	8	деятельность	1217	1.356
9	форма	330	1.339	9	работа	675	1.085	9	технология	1180	1.315
10	знание	328	1.331	10	обучение	669	1.076	10	время	1129	1.258
11	материал	321	1.303	11	компьютер	640	1.029	11	человек	1126	1.254
12	обучение	305	1.238	12	исследование	640	1.029	12	игрок	1020	1.136
13	исследование	282	1.145	13	зависимость	636	1.022	13	подросток	975	1.086
14	деятельность	265	1.076	14	образ	633	1.018	14	компьютер	917	1.022
15	образование	264	1.071	15	уровень	631	1.014	15	проблема	893	0.995
16	структура	261	1.059	16	технология	629	1.011	16	исследование	870	0.969
17	россия	260	1.055	17	группа	608	0.977	17	система	852	0.949
18	успех	256	1.039	18	система	590	0.949	18	образ	848	0.945
19	метод	255	1.035	19	результат	572	0.920	19	уровень	809	0.901
20	активность	246	0.998	20	проблема	559	0.899	20	возможность	792	0.882

Рис. 3. Терминограммы как результат абсолютного ЧОЗ по трем корзинам тематического массива текстов «Компьютерные игры». Запрос произведен по ЭБ «Humanitarian» Университета ИТМО 23.03.2016

Результатом относительного частотно ориентированного запроса является частотно-ранжированный список существительных, входящих только в те абзацы, которые содержат заданный пользователем термин (тем самым терминограмма строится относительно этого термина). Поиск может проводиться одновременно по 1, 2 или 3 корзинам ресурсов (рис. 4).

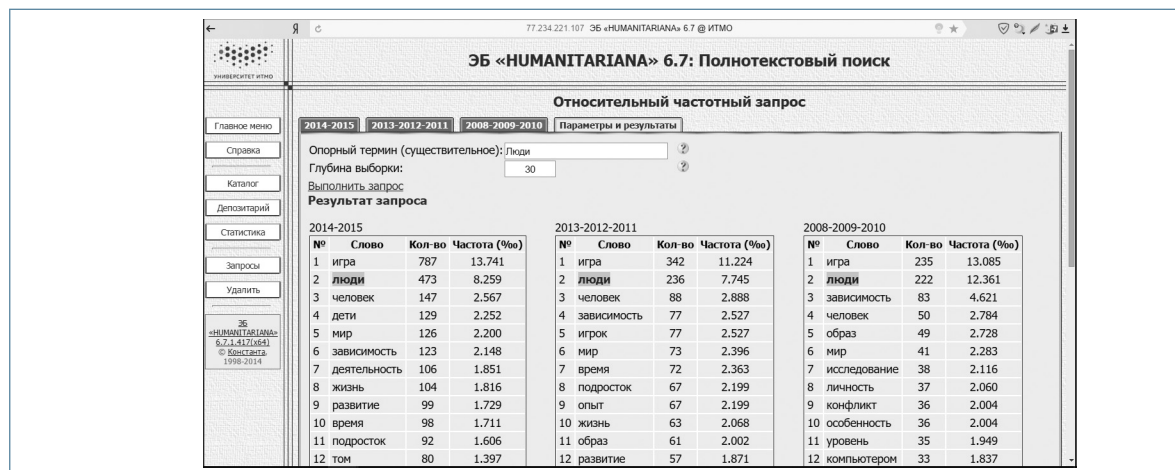


Рис. 4. Пример автоматического построения «терминограмм» (ЧОЗ относительно термина «люди») по трем корзинам ресурсов. Запрос произведен по ЭБ «Humanitarian» Университета ИТМО 23.03.2016

Все термины, входящие в терминограмму, являются активными. «Кликнув» любой из них, можно выйти на уже сформированный абзачно-ориентированный запрос по данному термину и эксплицировать его микроконтекст. Среди реализуемых с помощью контекстного поиска сервисов:

- текстологический анализ документа;
- выявление и описание предметной области документа;
- составление списка ключевых слов;
- сравнительный анализ предметных областей, разных авторов или разных документов;
- проверка текстов на плагиат и т.д.

Инструментарий ЭБ «Humanitarian» позволяет реализовать не только стандартные функции нахождения требуемых информационных ресурсов, их описания, каталогизации, хранения, поиска, присущие классическим автоматизированным информационным библиотечным системам, но и обеспечивает максимально возможную автоматизацию работы с тематическими коллекциями запросов.

В структуре ЭБ «Humanitarian» имеется подсистема «Личные ресурсы». Она предназначена для создания и управления совокупностью персональных хранилищ информации. Это могут быть:

- подборки библиографических описаний с прикрепленными к ним полнотекстовыми ресурсами;
- произвольные абзацы или их совокупности, полученные в результате полнотекстового поиска;



- терминограммы (частотно-ранжированные списки терминов);
- списки литературы, используемые при написании обзоров, научных отчетов и публикаций, цитировании источников;
- готовые структуры пользовательских запросов, тематических подборок материала.

Сервисы ЭБ «Humanitaria» могут быть использованы и для создания качественно нового вида информационных ресурсов — тематических коллекций полнотекстовых запросов. Запросы могут использоваться как готовое тематизированное (экспертное, контекстное) знание, расширяющее состав информационных ресурсов электронной библиотеки, и как пользовательский инструмент для создания и расширения коллекций.

Дополнительные возможности поиска и презентации результатов [6]:

- показ «кластера абзацев», в котором находится найденный абзац (открытие нескольких абзацев, расположенных в первичном документе до и после найденного);

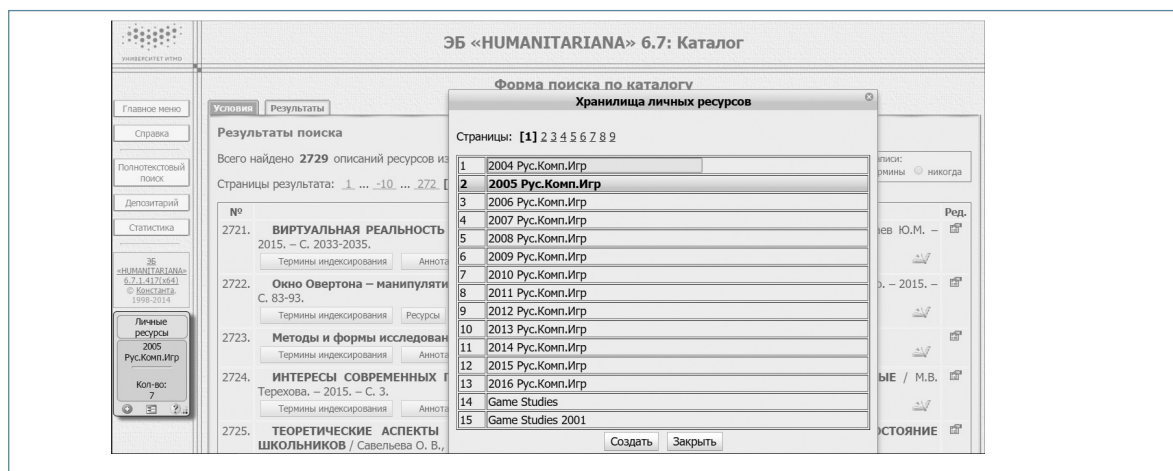


Рис. 5. Пример персональных хранилищ информационных ресурсов. Запрос произведен по ЭБ «Humanitaria» Университета ИТМО 23.03.2016

- отображение графического оригинала страницы текстового ресурса, содержащей найденный абзац текста;
- возможность сохранения запросов зарегистрированного пользователя для последующего формирования личного кабинета пользователя.

Сервисы ЭБ «Humanitaria» как инструмент исследования обеспечивают интерпретируемость и воспроизводимость полученных результатов, позволяют обосновать актуальность и практическую значимость тематики предполагаемого исследования.

Экспериментальной базой для разработки методики послужили тематические коллекции материалов по двум исследовательским тематикам — электронное правительство и игровой компьютерный опыт. Создание и апробация методики проводились в рамках подготовки магистерских диссертаций по направлению «Прикладная информатика» (Университет ИТМО, Факультет технологического менеджмента и инноваций) и поискового научного исследования (Социологический институт РАН, грант РФФИ 16-06-00368.а, 2016–2018 гг.), выполняемого с участием сотрудников университета ИТМО.

База данных ЭБ «Humanitarianiana», используемая для проведения исследования и апробации методики, была сформирована из трех информационных источников:

- архива научных публикаций по тематике «Электронное правительство и электронные услуги» (70 ресурсов состоящих из 5,2 тыс. новостных сообщений) за 2011–2015 гг.;
- научной периодики, размещенной в открытом доступе на портале научной электронной библиотеки eLibrary (<http://elibrary.ru/>); при поиске на портале задавалось ключевое слово «компьютерные игры»; хронологические рамки поиска 2005–2015 гг., всего для анализа было подготовлено 604 файла;
- научных англоязычных журналов «Game Studies» (выходит с 2001 г.), «Games and Culture» (выходит с 2006 г.) [7], которые первыми начали формировать социально-гуманитарную теорию компьютерных игр.

Разработанная в ИТМО методика использования сервисов контекстного поиска и извлечения контекстного знания в организации НИР применима для проведения как студенческих, так и независимых научных исследований, подготовки преподавателями и научными сотрудниками университета аналитических и библиографических обзоров, научных публикаций. Методика предназначена для поддержки следующих этапов НИР (таблицы 1–3):

- 1 постановка проблемы;
- 2 выдвижение и обоснование темы НИР/ВКР;
- 3 теоретическое исследование;
- 4 экспериментальные исследования;
- 5 анализ и сопоставление результатов.

Постановка проблемы — этап для исследователей. Основной сервис, используемый на данном этапе, — формирование частотно-ранжированного списка терминов, характеризующего предметную область исследований (смысловой макроконтекст). Он позволяет сформулировать тематику научных исследований в предметной области, определить динамику интереса научных кругов к отдельно взятым темам и проблемам, выделить наиболее и наименее интересные аудитории ракурсы научной проблемы.

Состав работ	Сервисы	Сценарий и результаты действий пользователя
Выбор направления исследования (тематика НИР) Выявление требований заказчиков (при наличии)	Каталог Персональные хранилища информационных ресурсов	1. Контекстный поиск по каталогу ЭБ 2. Содержательное структурирование произвольной совокупности документов для формирования хранилищ информационных ресурсов по предметной области
	Полнотекстовый частотно-ориентированный контекстный поиск	1. Формирование частотно-ранжированного списка терминов, характеризующего предметную область исследований 2. Анализ частотной выборки релевантных терминов для составления списка ключевых слов

Табл. 1. 1-й этап НИР: постановка проблемы

Тематика НИР магистрантов «определяется направлениями исследований выпускающей кафедры и других организаций, принимающих совместно с вузом участие в исследованиях. Тема НИР выбирается студентом под руководством преподавателя с учетом личных наклонностей и интересов студента» [8].

Результатом работы магистранта с сервисами ЭБ «Humanitaria» по выбранной тематике исследования на данном этапе является:

- выбор и обоснование актуальности темы, способов решения задач исследования (концептов, выделенных экспликацией предметной области);
- утверждение темы и разработка плана НИР;
- составление библиографических указателей, списков источников, каталогов.

Основные итоги данного этапа — подготовка материала для аналитики (терминограммы), формирование личных тематических коллекций, подготовка аналитического обзора по теме исследования (на основе эксплицированного линейно упорядоченного списка «горизонтальных» микроконтекстов) с использованием контекстов для цитирования фрагментов текста.

Состав работ по этапам НИР	Сервисы	Сценарий и результаты действий пользователя
Сбор, изучение, анализ результатов фундаментальных и поисковых исследований Формулирование тем НИР/ВКР и их сравнительная оценка	Абзацно-ориентированный контекстный поиск Рубрикатор депозитария Архив запросов и результатов поиска	1. Поисковые запросы разного типа с целью извлечения контекстного знания на основе списков ключевых слов 2. Сохранение созданных запросов, результатов поиска 3. Анализ запросов
Выбор и обоснование актуальности темы, способов решения задач исследования Утверждение темы и разработка плана НИР Подбор научной литературы по теме и ее проработка	Комбинация частотно-ориентированных и абзацно-ориентированных видов контекстного поиска	1. Поисковые запросы в хранилищах информационных ресурсов 2. Анализ выделенных контекстов (горизонтальных и вертикальных) 3. Экпликация смысловых «вертикальных контекстов» и тематизация материала 4. Экпликация «горизонтальных контекстов» употребления терминов в рамках авторского абзаца
Составление отчета по НИР	Персональные хранилища информационных ресурсов	Создание личных тематических коллекций электронных документов Формирование библиографического списка описаний хранилищ ресурсов

Табл. 2. 2-й этап НИР: выдвижение и обоснование темы НИР

Состав работ	Сервисы	Сценарий действий пользователя
Составление аналитического обзора Подготовка рефератов, научных публикаций по теме НИР	Персональные хранилища информационных ресурсов Абзацно-ориентированный контекстный поиск	1. Создание личных тематических коллекций электронных документов 2. Эксплицирование «горизонтальных» микроконтекстов из тематически ориентированных абзацев различных документов 3. Использование сервисов поиска для цитирования работ
Составление списка терминов по НИР	Абсолютный частотно-ориентированный контекстный поиск	Формирование на основе результатов частотно-ориентированного запроса списка терминов
Составление отчета по НИР – аналогично таблице 2		

**Табл. 3.** 3-й и 4-й этапы НИР: теоретическое и экспериментальные исследования

На этапе анализа и сопоставления результатов может быть использована любая комбинация сервисов ЭБ «Humanitarianiana». Особенность сценария действий пользователя состоит в использовании сервисов полнотекстового поиска (эксплицированных контекстов) для цитирования работ при подготовке научных публикаций и отчетов.

\* \* \*

Использование сервисов открытой распределенной библиотечной среды «Humanitarianiana» позволяет повысить управляемость и интенсифицировать проведение научных исследований, в первую очередь аналитического характера, а также создать собственные тематические хранилища и осуществлять доступ к тематическим хранилищам и коллекциям других организаций.

Апробация методики показала, что возможности и сервисы контекстного поиска ЭБ «Humanitarianiana» позволяют:

- более оперативно организовать предварительный этап НИР, связанный с рассмотрением научной тематики, выбором научного руководителя и темы исследования;
- сократить трудоемкость обработки информационных источников и ресурсов.

Проведенная работа намечает пути дальнейшего совершенствования организации научных исследований, аналитической и проектной деятельности с помощью уже реализованных и разрабатываемых сервисов ЭБ «Humanitarianiana».

ЛИТЕРАТУРА

1. **Демин И. С. Поиск научной и учебной информации в сети Интернет** // Вестник Тамбовского университета. Серия: Гуманитарные науки. — 2008. — Вып. 9. — С. 446–450.
2. **Голанова А. В. Некоторые аспекты обучения технологии поиска информации в сети Internet** / А. В. Голанова, Е. И. Голикова // Царскосельские чтения. — 2012. — Вып. 16. — Т. 4. — С. 129–134.
3. **Аюшеева Н. Н. Исследование и разработка моделей и методов поиска информационных образовательных ресурсов в электронной библиотеке:** автореф. дис. ... канд. техн. наук: 05.13.11. — Улан-Удэ, 2004.
4. **Шуклин А. Кому нужен контекстный поиск?** URL: <http://www.cnews.ru/reviews/free/Bl2012/articles/articles2.shtml> (дата обращения 15.03.2016).
5. **О модельном законе «О международном информационном обмене»: постановление Межпарламентской Ассамблеи государств — участников Содружества Независимых Государств от 26 марта 2002 г.** № 19–7. URL: <http://base.consultant.ru/cons/cgi/online.cgi?req=doc; base=INT; n=13288> (дата обращения 27.03.2016).
6. **Ляпин С. Х. Тематические коллекции полнотекстовых запросов для изучения контекстного знания (проект Humanitarianiana)** / С. Х. Ляпин, А. В. Куковякин // Сборник научных трудов XVIII Объединенной научной конференции «Интернет и современное общество», Санкт-Петербург, 23–25 июня 2015 г. — СПб: Университет ИТМО, 2015. — С. 216–224.
7. **Кононова О. В. Игровой компьютерный опыт как тема научной периодики в фокусе автоматизированного контекстного поиска** / О. В. Кононова, О. В. Сергеева, Е. А. Крутько, Е. А. Орех. // Материалы IV Международной научно-практической конференции «Коммуникации в социально-гуманитарном знании, экономике, образовании». Минск, 7–9 апреля 2016 г. — Минск: Издательство БГУ, 2016. — С. 251–253.
8. **Рекомендации по организации проведения научно-исследовательской работы студента в СПбГУ ИТМО от 2009 г.** URL: <http://uop.ifmo.ru/file/stat/17/nirstudenta.pdf> (дата обращения 15.03.2016).