Карманова Анастасия Александровна, Табаков Константин Андреевич

СИСТЕМА ПОСТРОЕНИЯ РЕКОМЕНДАЦИЙ КОНФЕРЕНЦИЙ НА ОСНОВЕ АНАЛИЗА НАУЧНЫХ ИНТЕРЕСОВ ПОЛЬЗОВАТЕЛЯ

Данная работа посвящена описанию системы анализа конференций. Разработанная система автоматически определяет тематику конференций по относящимся к ним текстовым документам на основе разработанного нами алгоритма. Алгоритм базируется на модели "мешок слов", таксономии ACM CCS и учитывает иерархическую структуру словаря терминов. Разработанный алгоритм используется также при анализе научных интересов пользователей. Система строит индивидуальные рекомендации. Кроме того, система содержит вопросно-ответный модуль, который предоставляет пользователям возможность получить обратную связь.

Адрес статьи: www.gramota.net/materials/1/2017/6/11.html

Источник

Альманах современной науки и образования

Тамбов: Грамота, 2017. № 6 (119). С. 33-44. ISSN 1993-5552.

Адрес журнала: www.gramota.net/editions/1.html

Содержание данного номера журнала: www.gramota.net/materials/1/2017/6/

Толстой считал, что жизнь умерших людей не прекращается в этом мире. Если смерть, рассуждал Толстой, есть уничтожение, то сама жизнь отравлена страхом от этого, и чтобы хоть как-нибудь избавиться от него, он построил свою философию жизни и смерти. Смерть не есть зло, так как это есть неоспоримый закон Бога. Сущность смерти писатель представлял себе так: она является достижением такого состояния, при котором нет ни желаний, ни частного существования, а только слияние части с целым. «Никто не знает, что такое смерть, но все ее боятся, считая ее величайшим злом, хотя она может быть и величайшим благом» [5, с. 79].

Толстой верил в способности личности, ее саморазвитие и самосовершенствование. Учение Толстого – важный вклад в развитие гуманистической мысли. Это и защита гуманности, и выдвижение человеколюбия на первый план как главной ценности в жизни людей, основного принципа морали и общественных отношений. Это критика антигуманизма и концепция естественности гуманизма и добра и неестественности безнравственных отношений, войн, насилия, вражды. Гуманистические идеи Толстого вышли за границы времени и пространства.

Толстой наделен безграничной верой в человека, считает, что люди в силах сделать нравственно-волевое усилие над собой для воздержания от противных истине слов, мыслей, поступков.

Наследие гениального писателя особенно значимо в наши дни своей нравственной требовательностью, гуманистическим принципом, беспредельной верой в улучшение человека. Идеи Толстого и сегодня актуальны, они оказывают огромное влияние на нравственный мир человека, на то, как он решает для себя вопросы жизни, смерти и бессмертия.

Список источников

1. **Ли Со Ен.** Религиозно-философская антропология Л. Н. Толстого. М., 1996. 150 с.
2. **Попов Г. А.** Проблема жизни и смерти в религиозно-философской антропологии Толстого. М., 2006. 350 с.
3. **Толстой Л. Н.** Война и мир // Толстой Л. Н. Собрание сочинений: в 4-х т. М.: Эксмо, 2010.
4. **Толстой Л. Н.** О жизни. Изд. 2-е. М.: А. П. Никифоров, 1903. 288 с.
5. **Толстой Л. Н.** Путь жизни. М.: Эксмо, 2009. 448 с.
6. **Толстой Л. Н.** Смерть Ивана Ильича. М.: Худож. литература, 1987. 304 с.

LEO TOLSTOY'S UNDERSTANDING OF LIFE AND DEATH

Evsina Ekaterina Vital'evna
Lipetsk State Pedagogical P. Semenov-Tyan-Shansky University
ekaterina.evsina@bk.ru

This article is an examination of Leo Tolstoy's opinion about the problem of life and death. What is life for the writer, and what is death? The paper gives a philosophical description of Tolstoy's ideas of death on the basis of the analysis of his works ("War and Peace", "The Death of Ivan Ilyich") and carries out a study of Leo Tolstoy's humanistic and moral ideas and their significance in the life of the modern man.

Key words and phrases: morals; good; life; death; fear of death; God; humanism.

УДК 004.4
Технические науки

Данная работа посвящена описанию системы анализа конференций. Разработанная система автоматически определяет тематику конференций по относящимся к ним текстовым документам на основе разработанного нами алгоритма. Алгоритм базируется на модели «мешок слов», таксономии ACM CCS и учитывает иерархическую структуру словаря терминов. Разработанный алгоритм используется также при анализе научных интересов пользователей. Система строит индивидуальные рекомендации. Кроме того, система содержит вопросно-ответный модуль, который предоставляет пользователям возможность получить обратную связь.

Ключевые слова и фразы: тематическое моделирование; интеллектуальный анализ текста; рекомендательная система; таксономия; «мешок слов»; вопросно-ответные системы.

Карманова Анастасия Александровна
Табаков Константин Андреевич
Новосибирский государственный университет
anast.karmy.aa@gmail.com; konkov90@gmail.com

СИСТЕМА ПОСТРОЕНИЯ РЕКОМЕНДАЦИЙ КОНФЕРЕНЦИЙ НА ОСНОВЕ АНАЛИЗА НАУЧНЫХ ИНТЕРЕСОВ ПОЛЬЗОВАТЕЛЯ

Введение

В современном мире объемы доступной информации растут с беспрецедентной скоростью, и к началу 2017 года число проиндексированных web-страниц стремится к 4,5 миллиардам [12]. Следовательно,

невозможно не только изучить все цифровые документы, но и разобраться со всей существующей информацией в рамках одной тематики. Поэтому для исследователей в какой-либо области критически необходима организация исследовательского сообщества для совместной работы. Коммуникация между исследователями, учеными может происходить посредством публикации статей, участия в конференциях, симпозиумах и т.д.

Из всего вышесказанного можно сделать вывод, что если молодой ученый делает свои первые исследовательские шаги, ему следует участвовать в различных мероприятиях научной направленности по его теме, чтобы присоединиться к исследовательскому сообществу. Необходимо отметить, что участие в таких мероприятиях помогает исследователю познакомиться с контекстом, в котором он выполняет свою работу, а значит, выяснить, какие пути решения поставленной им проблемы существуют в сообществе. Кроме того, участие в конференциях, семинарах, симпозиумах – это отличная возможность поделиться своими идеями и «опробовать» их на аудитории. В результате такого «тестирования идей» можно получить ценные рекомендации по исправлению ошибок и недочетов. Мнение компетентных специалистов в этом случае может оказаться очень полезным.

Однако на сегодняшний день проходит такое огромное количество конференций и других мероприятий по различным направлениям исследований, что разобраться в них непросто. Следующей проблемой является то, что у многих из этих мероприятий тематика указана широко, и непонятно, затронет ли та или иная конференция интересующие конкретного исследователя темы. Поэтому цель данной работы – помочь молодому исследователю в поисках интересных для него мероприятий научного характера.

Цель разработки

Итак, цель нашей работы – помочь молодым ученым оставаться в курсе исследований по темам, над которыми они работают. Чтобы знать, что происходит в научном сообществе, нужно с ним активно взаимодействовать – участвовать в различных мероприятиях, делиться своими достижениями и анализировать идеи других исследователей. Но из всего многообразия конференций, симпозиумов и семинаров бывает сложно выбрать действительно полезные для работы. Посетить их все, очевидно, невозможно. Поэтому решено было создать информационную систему, которая бы сопровождала молодых ученых в процессе выбора мероприятий.

Пользователь такой системы – исследователь, который является новичком в научном сообществе и нуждается в совете, или более опытный ученый, который уже участвовал в ряде конференций, но хочет расширить кругозор в своей предметной области. Система должна предоставить возможность узнать интересующую пользователя информацию о конференциях, а также получить индивидуальную рекомендацию тех конференций, которые могут быть для него полезны.

Для своего исследования мы остановились на предметной области «Информационные технологии». Эта тема, с одной стороны, хорошо нам знакома, а, с другой стороны, широко и многогранно освещена на различных конференциях и в печати. Из-за огромного количества мероприятий в этой предметной области проблема выбора интересных для исследования событий становится особенно острой.

Помимо этого, мы сузили область нашей работы до поиска только конференций, для простоты проведения исследования. В дальнейшем планируется рассматривать также и другие способы коммуникации внутри научного сообщества.

Задачи системы

Таким образом, перед нами стояла цель разработать систему, которая обрабатывает ИТ-конференции и на основании этой информации может отвечать на вопросы о мероприятиях и строить индивидуальные рекомендации для пользователей. Достижение этой цели разбивается на несколько задач.

Во-первых, нам нужна модель предметной области, чтобы определить концептуальные границы разрабатываемой системы.

Во-вторых, нам нужно уметь определять научные интересы пользователя, чтобы работа системы была актуальной для него.

В-третьих, нам нужны источники информации о конференциях, которые обеспечат систему достаточными первоначальными сведениями для последующего анализа. Такими источниками могут быть сайты самих конференций или готовые ресурсы-агрегаторы.

В-четвертых, нам нужно уметь понимать, какие темы были затронуты на конференции. Здесь возникает следующая проблема: чаще всего, доступная информация о конференции затрагивает организационные моменты (время проведения, место проведения, сборы и т.д.), а о темах, обсуждаемых на конференции, известно не так много: общая тематика и, возможно, название подсекций. Этого может быть недостаточно, ведь предметная область, с которой мы работаем, «Информационные технологии», очень обширна, и предугадать, что конкретно будет обсуждаться, оказывается непросто.

В-пятых, необходимо уметь строить рекомендации конференций, которые были бы основаны на научных интересах исследователя и могли бы помочь пользователю продвинуться в его работе.

И, наконец, пользователь должен быть обеспечен возможностью обратной связи с системой. Ему должен быть предоставлен удобный способ узнать всю интересующую его информацию о мероприятиях, не выходя из системы. Эта возможность в нашей системе реализуется посредством вопросно-ответного модуля.

1. Построение модели предметной области

Для того чтобы суметь помочь исследователю в области «Информационных технологий» с выбором мероприятий, система должна понимать, что ему интересно, а следовательно, говорить с ним на одном языке. Имеется в виду, что профессиональная терминология, используемая пользователем, должна быть понятна системе.

Таким образом, возникает необходимость в разработке концептуальной модели предметной области. Эта задача требует вовлечения специалистов в самых различных областях ИТ и серьезной проработки. Поэтому было решено использовать готовые модели предметной области, надежность которых доказало многократное использование в различных научных работах.

При выборе модели основная задача, которую мы ставили перед собой, – понимать профессиональную терминологию, так как лексика позволит определить основной вектор исследования. Следовательно, было решено найти таксономию, достаточно полно описывающую ИТ.

Первоначально была рассмотрена Универсальная десятичная классификация. Классификационная система УДК – это система классификационных таблиц, наглядно представляющих различные тематические и аспектные части системы с той или иной степенью подробности [2]. Эта система используется во всем мире для научных работ, произведений литературы, организации картотек и хранилищ данных. Каждой теме приписан код, иерархии тем соответствует иерархия кодов.

Эта классификация содержит и отрасль «Информационные технологии» (соответственный код «004»), с которой мы работаем. Для этой отрасли классификация содержит более 800 терминов, что является преимуществом. Однако данная система содержит ряд недостатков.

Во-первых, УДК описывает «Информационные технологии» достаточно верхнеуровнево. Во-вторых, классификация не часто пополняется и является неактуальной. При этом отрасль «ИТ» быстро растет и развивается, и важно, чтобы классификация терминов оставалась актуальной. Наконец, эта система охватывает все области знания, и потому каждая отдельная узкая область оказывается неверифицированной узким экспертом.

Аналогичными достоинствами и недостатками обладает Государственный рубрикатор научно-технической информации (ГРНТИ).

Помимо УДК и ГРНТИ, была рассмотрена классификация ACM *Computing Classification System* [10], которая традиционно считается «золотым стандартом» для области «Информационные технологии». Эта система достаточно хорошо описывает предметную область нашей работы: в ACM содержится более 2000 терминов из области ИТ; эти данные могут быть представлены в виде дерева глубиной не более 6 уровней. Кроме того, данная система классификации является публичной и доступной. Наконец, важным фактором является актуальность базы, потому что область «Информационные технологии» – быстро развивающаяся, постоянно появляются новые термины.

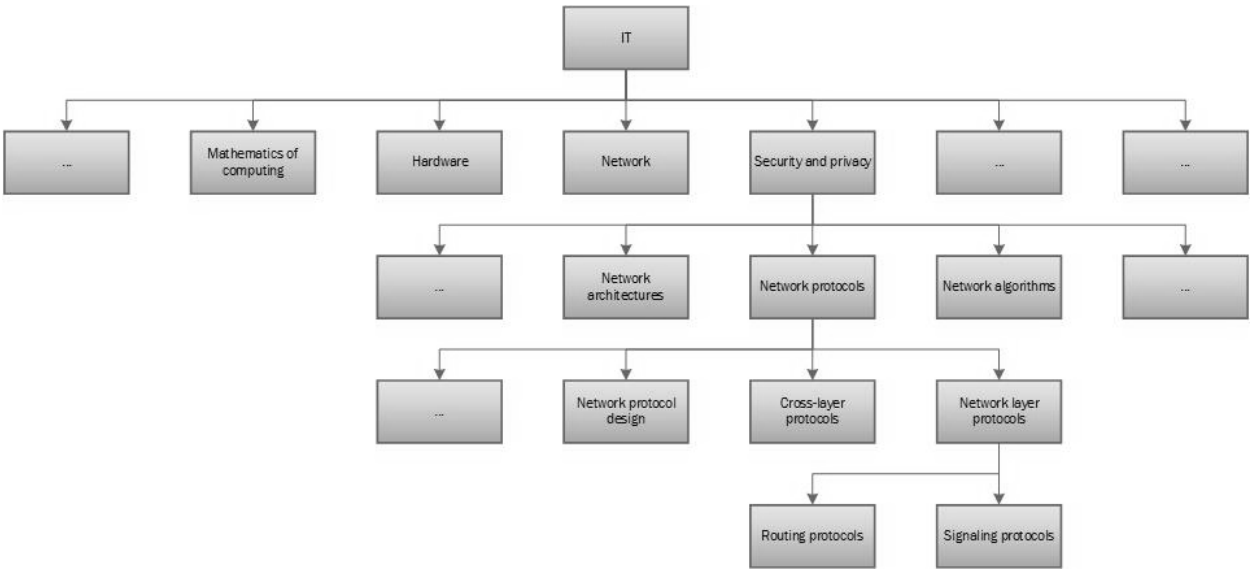

Рис. 1. *Классификация ACM*

Учитывая все вышеперечисленные преимущества ACM CCS, эта система была выбрана в качестве модели предметной области. Однако ACM CCS полностью англоязычная и требует локализации для работы с русскоязычными конференциями. Данная проблема решается с помощью полуавтоматизированного перевода.

1.1. Локализация таксономии ACM CSS

Локализация таксономии проходит в два этапа. На первом этапе дерево ACM CCS переводится поверхностно с помощью ресурса *Google Translate*. Данный ресурс был выбран нами по ряду причин. Во-первых, он является доступным и обладает открытым API. Во-вторых, ресурс является бесплатным. В-третьих, ресурс является надежным источником перевода, так как с момента выпуска сервиса в 2006 году огромное количество людей доверило свои переводы *Google Translate*. Действительно, по данным самого сервиса, количество пользователей превысило 500 миллионов и продолжает расти.

Однако сервис *Google Translate* не является узкоспециализированным в области ИТ, поэтому полностью полагаться на него нельзя. В связи с этим, было принято решение проводить дополнительную верификацию и настройку экспертом.

На втором этапе локализации эксперт в предметной области может редактировать дерево предметной области. Для этих целей в системе есть модуль «Настройка предметной модели», доступ к которому можно получить только в режиме администратора. Таким образом, рядовой пользователь разрабатываемой нами системы не имеет доступа к внутренним настройкам алгоритмов работы.

В модуле «Настройка предметной модели» работа эксперта выглядит следующим образом. В качестве исходной модели, как уже было сказано ранее, используется граф классификации ACM, однако при желании эксперт может импортировать в систему другую модель. Данные модели визуализируются, каждой вершине дерева соответствует термин на английском языке. Эксперт выбирает опцию автоматического перевода с помощью *Google Translate*, в результате чего каждому узлу графа сопоставляется набор вариантов перевода на русский язык термина на английском языке. Далее эксперт может вносить изменения в переводы, добавлять и удалять варианты. Кроме того, эксперту доступна возможность редактировать сам граф: добавлять и удалять узлы.

После того, как все изменения были внесены, эксперт сохраняет модель, и отредактированный вариант будет использоваться для последующих вычислений. Дополнительно эксперт может экспортировать модель.

Предметная область «Информационные технологии» быстро развивается, сегодня проходит огромное количество мероприятий на эту тему. Соответственно, исследователю в данной сфере знаний бывает непросто выбрать действительно подходящие конференции из всего многообразия.

Поэтому важной задачей нашей системы является возможность определять научные интересы пользователя, чтобы быть действительно актуальной для него. Определение полезных для молодого ученого тем происходит на основе информации, которой он делится с системой с последующей обработкой и анализом.

Пользователи регистрируются в системе, и источником информации о них является исключительно система. При регистрации пользователь должен ответить на несколько вопросов – заполнить свой профиль. Понятно, что пользователю не хочется тратить большое количество времени на заполнение длинных анкет, поэтому был выбран подход, при котором последующие вопросы задаются на основе предыдущих ответов. Итак, заполнение информации профиля происходит в несколько этапов:

1. Пользователь вводит базовые характеристики: логин, место работы, место учебы, должность.
2. На втором этапе пытаемся определить научные интересы пользователя. Пользователь может:
   a. Ввести ключевые слова, описывающие область его исследований.
   b. Загрузить документы, относящиеся к теме его исследований.
   c. Выбрать уже посещенные мероприятия.
3. На основе этой информации пользователю задаются вопросы о других мероприятиях.

Поддержание актуальной информации обеспечивается периодическими запросами системы на обновление информации. Пользователю достаточно ответить на несколько вопросов.

Итак, пользователь может помочь системе получать информацию несколькими путями.

Во-первых, пользователь может прямо указать ключевые для его исследования термины. Во-вторых, пользователь может указать, в каких конференциях он принимал участие. И, наконец, пользователь может загрузить в систему текстовые документы, касающиеся его исследования. К таким документам относятся аннотации, тезисы, статьи.

Разработанный алгоритм анализа научных интересов пользователя работает со всеми этими источниками как с текстовыми документами:

1. Каждое указанное ключевое слово рассматривается как текстовый документ, состоящий из одного слова – термина, который пользователь сообщил системе.
2. Текстовые документы, которые пользователь загрузил в систему.
3. Работа с конференциями также представляется как работа с текстовыми документами, которые относятся к той или иной конференции. Подробнее об этом – в параграфе «Определение тематики конференции».

1.2. Подходы к представлению текстов

Существует несколько подходов к работе с текстовыми документами. В рамках подхода «мешок слов» (“bag of words”) учитывается только, встречается ли слово в документе и, если да, то как часто. При этом игнорируются смысловые связи между словами, положение слов относительно друг друга в предложении, а также связи между самими текстами (в виде цитирования). Данная модель была предложена в 1975 году Дж. Солтоном [7] и в настоящее время является одной из самых распространённых в самых различных областях лингвистических исследований.

Подход «мешок термов» (“bag of terms”) является обобщением подхода “bag of words”. Под термом понимается символьное выражение объекта формальной модели, в данном случае языка. Поэтому для модели «мешок термов» в рассмотрение попадают не только слова, но и другие символьные выражения, такие как знаки препинания.

Векторная модель представления текста сопоставляет каждому тексту вектор из пространства, единого для всего корпуса текстов. Размерность этого пространства совпадает с количеством термов из формальной модели. Векторы, таким образом, представляют собой множество всех символьных выражений, а значение вектора для каждого текста задается частотой вхождения этого символьного выражения в текст.

Во всех описанных выше методах игнорируются порядок слов в предложении и тексте, связи между текстами в виде цитирования, возможная метаинформация о тексте (метки, теги). В нашем случае работа с моделями, учитывающими эти параметры, является неоправданным усложнением, поскольку объем корпуса текстов достаточно велик, тексты являются разрозненными и лишенными какой-либо метаинформации.

В случае с конференциями ИТ-тематики мы используем готовую таксономию терминов предметной области, поэтому множество T в данном случае будет полностью совпадать с множеством терминов из выбранной таксономии. А раз мы не учитываем порядок слов, то получаем, что каждый текст является «мешком термов», в котором учитывается только вхождение терминов предметной области.

1.3. Описание разработанного алгоритма

Для разработанного алгоритма пользователь представляется набором документов, связанных с его исследовательской деятельностью. Пусть для некоторого пользователя существует набор документов $D = \{d_1, d_2, \ldots, d_M\}$.

На первом этапе обработки всех текстовых документов из D каждый из этих документов представляется «мешком термов», в котором термами являются термины предметной области из локализованной версии таксономии ACM CCS. Обозначим множество терминов через $Term = \{term_1, \ldots, term_T\}$.

В рамках программного комплекса «RiskPanel» [1], для предметной области информационной безопасности была разработана вопросно-ответная система «QA-RiskPanel» [2]. Данная система основана на прецедентном подходе к моделированию предметных областей; база знаний при данном подходе моделируется в виде обобщенной нечеткой модели. Так как знания, представленные в базе, могут быть неполными или неточными, ответы на поставленные вопросы будут носить вероятностный характер. Для вопросов, которые могут быть заданы такой системе, была разработана классификация, состоящая из нескольких вопросных типов; возможность получать ответы на составные вопросы реализуется посредством основных операций классической логики.

В общем случае внутреннюю логику вопроса действительно можно описать

Рис. 2. *Выделение терминов предметной области в тексте*

В качестве частотной характеристики используем TF, так как она является достаточно простой для вычисления и при этом обеспечивает приемлемый результат.

Введем обозначение: частоту вхождения j-го термина в m-ый документ i-ого пользователя обозначим через $f_{m(i)j}$.

Далее «мешки термов», соответствующие различным документам одного пользователя, обобщаются в один. Это происходит путем сложения частот для соответствующих термов в различных документах. После чего результирующий «мешок термов» нормируется на интервале [0; 1]. В результате этих действий значение частоты вхождения термина $term_j$ в тексты i-ого пользователя будет вычисляться как:

$$F_{ji} = \frac{\sum_{m(i)=1}^{M(i)} f_{m(i)j}}{M(i)},$$

где M(i) – количество документов, соответствующих i-ому пользователю.

Итак, мы получили вероятности того, что тот или иной термин является актуальным для работы пользователя. Эти данные получены без учета отношений между терминами. Такие отношения обусловлены структурой таксономии предметной области.

ACM CCS, как было сказано ранее, имеет древовидную структуру, в которой каждому узлу дерева сопоставляется термин предметной области. Корневым узлом является самый общий термин, и чем ниже располагается термин в дереве, тем более узко термин описывает тематику.

На втором этапе алгоритма, после того, как был получен результирующий «мешок слов», происходит учет древовидной структуры словаря терминов.

Результатом работы второго этапа алгоритма является сопоставление каждому пользователю системы дерева терминов предметной области, размеченного вероятностями на интервале [0; 1]. Эти вероятности следует интерпретировать так: термин $term_j$ отвечает интересам i-ого пользователя с вероятностью w_{ji}.

С этой целью каждой конференции c_i было сопоставлено дерево терминов, размеченное таким образом, что каждому узлу-термину $term_j$ соответствует его первичная вероятность f_{ji}, полученная на первом этапе алгоритма. Сопоставление происходит в следующем порядке.

Построение дерева:

1. Сначала для i-го пользователя строим дерево, структура которого совпадает со структурой дерева предметной области. Каждому узлу этого дерева сопоставляем значение вероятности w_{ji}. На этом шаге $w_{ji} = 0$.

2. Берем обобщенный «мешок термов», полученный на первом этапе, в котором каждому термутермину $term_j$ соответствует вероятность его актуальности F_{ji} для пользователя i. Термины из «мешка термов» соответствуют узлам дерева предметной области. Присваиваем соответственно:

$$w_{ji} \leftarrow w_{ji} + F_{ji}.$$

3. Теперь для каждого узла j, начиная с листьев и двигаясь к корневому элементу, выполняем следующее: к вероятностям всех узлов, лежащих выше по дереву, чем узел j, прибавляем значение w_{ji}.

Таким образом мы получили дерево, каждому узлу которого соответствует вероятность того, что термин, содержащийся в этом узле, является актуальным для исследования пользователя.

Данный алгоритм может быть описан при помощи следующего псевдокода.

Вход:

- корпус текстовых документов $D = \{d_1, \ldots, d_M\}$, каждый из которых состоит из слов $word_1, \ldots, word_{S(m)}$, где $m \in [1;\ M]$;
- количество документов в корпусе M;
- набор терминов предметной области $Term = \{term_1, \ldots, term_T\}$;
- количество терминов K;
- частоты встречаемости терминов в документах $f_{11}, \ldots, f_{T1}, f_{1M}, \ldots, f_{TM}$;
- вероятности актуальности терминов $F_1, \ldots, F_K$;
- граф предметной области G (V, E) с вершинами $v_1, \ldots, v_k$, взвешенными функцией $weight(v_j)$.

Выход:

- граф предметной области G (V, E) с вершинами $v_1, \ldots, v_k$, взвешенными функцией $weight(v_j)$.

Алгоритм:

Шаг 1: Подсчитываем частоты встречаемости терминов и нормируем по TF для каждого документа:

While 1 ≤ m ≤ M
  While 1 ≤ s ≤ m
    While 1 ≤ j ≤ T
if (term[j] = word[s]) then f[m][j]++;
While 1 ≤ m ≤ M
  While 1 ≤ j ≤ T
    f[m][j] = TF (f[m][j]);

Шаг 2: Обобщаем данные по разным документам в один и нормируем по TF:

While 1 ≤ m ≤ M
  While 1 ≤ j ≤ T
    F[j] += f[m][j];
While 1 ≤ j ≤ T
  F[j] = **TF**(F[j]);

Шаг 3: Производим первичное заполнение графа

While 1 ≤ j ≤ T
  weight (v[j]) = F[j];

Шаг 4: Дополнительная разметка графа с учетом структуры:

While j1 ∈ [1; T] **move** j1 **from leaf to root**
While 1 ≤ j2 ≤ T
if v[j1] **is_more_general_than** v[j2] then v[j2] +=v[j1].

1.4. Пример работы алгоритма

Рассмотрим работу алгоритма более подробно на примере. На Рис. 3 изображен фрагмент таксономии терминов.

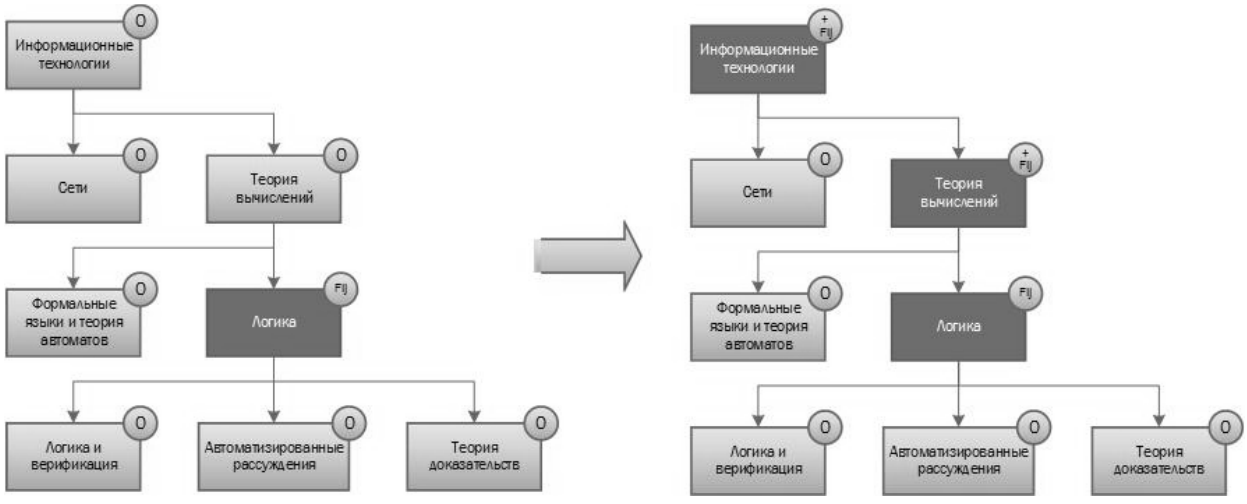

Рис. 3. *Разметка дерева с учетом структуры предметной области*

Пусть нам известно, что термин «Логика» актуален для пользователя с вероятностью f_{ij}. Тогда $w_{ij} = f_{ij}$ по пункту 2 алгоритма. Если термин «Логика» актуален с некоторой вероятностью, то справедливо, что и тема «Теория вычислений» относится к интересам пользователя с не меньшей вероятностью, что и более общая тема. Поэтому темам выше по дереву присваиваем то же значение w_{ji}.

Термин «Автоматизированные рассуждения» является более узким термином, чем «Логика», поэтому если пользователю интересна тема «Логика», то мы не можем утверждать, что и более узкая тема «Автоматизированные рассуждения» заинтересует пользователя.

1.5. Сбор и обработка информации о конференциях

Информацию о конференциях можно получать различными способами. Рекламные записи, записи в социальных сетях, почтовые рассылки и твердые носители – это источники, которые, безусловно, могут информировать людей о конференциях, но такие источники трудны в обработке и могут оказаться ненадежными. Самыми достоверными и удобными для работы источниками мы считаем сайты конференций и сайты-агрегаторы конференций.

Стоит заметить, что невозможно иметь ссылки на сайты всех существующих конференций по ИТ, поэтому в качестве первичного источника было решено использовать сайт-агрегатор, а более подробную информацию о конференции можно брать с ее сайта, если ссылка указана на сайте-агрегаторе.

Соответственно, перед нами стояла задача выбора ресурса информации о конференциях.

1.6. Обзор ресурсов информации о конференциях

Подробно рассмотрим все типы агрегаторов информации о конференциях. Условно их можно разделить на 3 типа.

1. *Порталы организаторов конференций.* Такие ресурсы помогают в информационной организации конференции. В задачи таких ресурсов обычно входят: подготовка web-сайта конференции, регистрация заявок участников, настраиваемая автоматическая проверка заявок, хранение материалов конференций. Отличительной особенностью является то, что они хранят информацию только об организованных ими мероприятиях. Примером таких ресурсов может служить информационная система «Конференции» [5].
2. *Специализированные ресурсы.* Такие источники хранят информацию только об определенных областях науки. Так, например, сайт «Грамота.ру», имеющий специализацию «русский язык и литература», содержит раздел о конференциях этой тематики. Зачастую такие ресурсы являются разделом сайтов институтов.
3. *Агрегаторы* – это ресурсы, которые наполняются автоматически или вручную, после чего они проходят процедуру верификации человеком и публикуются для общего доступа. Такие ресурсы содержат конференции различной тематики. Примером агрегаторов может служить портал «Конференции.ру».

Наша система нацелена на автоматическую обработку конференций, что накладывает ряд ограничений на источники информации:

1. *Открытость.* Закрытые источники информации нас не интересуют ввиду сложности извлечения данных из них.
2. *Объем источника.* Количество конференций, содержащихся в базе данных, должно быть достаточным для построения модели.
3. *Количество информации о конференции.*
4. *Пополняемость базы.* Информация о конференциях быстро устаревает, поэтому необходимо выбрать источник, содержащий актуальную информацию.
5. *Возможность автоматического извлечения информации о конференции из источника.*

Для первого приближения был выбран каталог конференций «Конференции.ру». Он находится в открытом доступе и содержит информацию об очных и заочных конференциях. На текущий момент в источнике зарегистрировано более 23000 конференций. Агрегатор содержит следующую информацию:

1. Место проведения.
2. Форма участия.
3. Язык.
4. Последний день подачи заявок.
5. Список организаторов.

2. Определение тематики конференции

При анализе конференций самой важной характеристикой является тематика, ведь на основании нее потенциальные участники решают, насколько им интересно участвовать в той или иной конференции. Однако определение тематики зачастую является проблемой. В самом деле, часто тематика конференций определена очень широко: «Гуманитарные науки», «Информационные технологии», «Медицина». Однако эти широкие тематики не дают специалисту в области оценить степень полезности конференции для конкретного научного исследования. Конференция по «Информационным технологиям» может рассматривать проблемы машинного обучения, искусственного интеллекта и написания технической документации, а может не касаться этих областей. Таким образом, возникает вопрос, как определить более точную тематику конференции, если ее описание содержит только верхнеуровневые темы обсуждения.

Решение состоит в следующем. Каждая конференция сопровождается набором текстовых документов: аннотация, тезисы, программа. Предлагается анализировать документы, описывающие конференцию, на основании терминов предметной области, встречающихся в них.

2.1. Основные подходы к тематическому моделированию

Возможность автоматического определения темы документов является критически важной в связи с огромными объемами доступной цифровой информации и невозможностью обработать эти объемы вручную.

Построение тематических моделей используется в различных сферах для классификации, кластеризации, ранжирования и аннотирования различных источников: новостей, записей в блогах, статей, научных работ. На основе этих знаний можно определять актуальные темы в той или иной предметной области, анализировать авторов текстовых документов, строить рекомендательные модели.

Тематическое моделирование – способ построения модели коллекции текстовых документов, которая определяет, к каким темам относится каждый из документов [4]. Для такой задачи, как тематическое моделирование, входными данными являются корпус текстов и словарь терминов предметной области.

Результатом тематического моделирования является сопоставление каждого документа одной или нескольким темам предметной области.

2.1.1. Базовые тематические модели

Существует огромное количество тематических моделей, но базовые из них, как правило, работают с текстами как с «мешком слов», такая модель представления текста является самой распространенной и простой в работе.

Нами было рассмотрено несколько тематических моделей.

Вероятностный латентный семантический анализ (Probabilistic Latent Semantic Analysis, PLSA) – модель, которая на вход получает корпус текстов D = {d1, d2, …,d_N}, термины предметной области T = {t1 t2, …, t_M} и количество тем K. В ходе работы алгоритма определяется скрытый, или латентный, слой переменных Z = {z1, z2, …, z_K} – темы документов. В результате, каждый документ описывается вектором, в котором каждая i-ая компонента описывает долю zi-ой темы в документе. Однако вероятностная модель не описывает ни закон распределения этих долей, ни вероятности самих документов.

Эти ограничения были разрешены в модели латентного размещения Дирихле (Latent Dirichlet Allocation, LDA). В LDA предполагается, что каждое слово в документе порождено некоторой латентной темой, при этом в явном виде моделируется распределение слов в каждой теме, а также априорное распределение тем в документе.

Существует ряд расширений LDA. Автор-тематическая модель (Author-Topic Model) [11] позволяет анализировать тексты и устанавливать взаимосвязи между текстовыми документами и авторами. Возможность учитывать зависимости между словами документа описывается в модели HMM-LDA, основанной на скрытых марковских моделях (Hidden Markov Model, HMM).

В нашем случае конференции представлены наборами документов, для которых связи между словами в документе и связи между самими документами не определены. Документы описываются моделью «мешок слов» на пространстве терминов из таксономии предметной области.

Представленные выше алгоритмы требуют, чтобы количество тем для каждого документа было задано заранее, однако, в нашем случае мы не можем заранее дать оценку этому количеству для каждого документа. Кроме того, эти подходы не учитывают иерархическую структуру таксономии терминов.

В связи с этим был разработан собственный подход к определению тем документов и, как следствие, конференций.

2.2. Определение тематики конференций

В нашем случае, задача определения темы конференции является аналогичной задаче выявления научных интересов пользователей, потому что в основе обеих задач лежит общая проблема – определение тем, о которых говорится в корпусе текстовых документов.

Поэтому для каждой конференции по набору документов, сопровождающих ее (программа, аннотация, тезисы), строится взвешенное дерево, аналогичное тому, которое мы строили при анализе интересов пользователя. Каждый узел этого дерева – термин предметной области, каждому термину приписано числовое значение из интервала [0; 1], которое соответствует вероятности того, что конференция посвящена этой теме.

Будем считать, что темами конференции являются все темы, значение частоты для которых выше порога *Threshold*, который настраивается вручную.

3. Построение индивидуальных рекомендаций

Рекомендательные системы – это системы, которые анализируют интересы пользователей и на основании полученных данных пытаются представить наиболее интересные для пользователя предложения. В настоящее время существует множество рекомендательных систем, предлагающих подбор различных товаров и услуг. Рекомендательные механизмы сортируют огромные объемы данных для выявления потенциальных предпочтений пользователей.

3.1. Обзор существующих рекомендательных систем

Amazon – один из лидеров области. *Amazon* рекомендует пользователям книги и другие товары, основываясь на том, что они покупали и просматривали ранее, какие оценивали товары и какие оставляли отзывы.

Другим примером рекомендательных систем могут служить сервисы *Last.fm* и *Pandora*, рекомендующие музыку. При построении рекомендации *Last.fm* использует не только рейтинги других пользователей, но и «внешние» данные о музыке – автор, стиль, дата, тэги. Сервис *Pandora* оценивает «содержание» музыкальной композиции, используя экспертную оценку: профессиональные музыканты анализируют композицию по нескольким сотням атрибутов через сервис *Music Genome Project*.

3.2. Фильтрация информации

Существует большое количество типов рекомендательных систем, рассмотрим два основных из них.

Контентная фильтрация формирует рекомендацию на основе предыдущего поведения пользователя. Пользователю рекомендуются объекты, похожие на те, с которыми этот пользователь уже взаимодействовал. Сходство оценивается по содержимому объектов. Надо заметить, что существует сильная зависимость от предметной области, и польза рекомендаций в некоторых случаях может быть ограничена.

Системы коллаборативной фильтрации – это модели, которые пытаются предсказать, насколько пользователю понравится тот или иной продукт, получая на вход данные о том, как он и другие пользователи оценивали этот и другие продукты в прошлом. Коллаборативная фильтрация – это самый популярный ныне вид рекомендательных систем. Точность рекомендации в таких системах зависит от объемов накопленной статистики.

Для рекомендации в таких системах используется история оценок не только самого пользователя, но и других пользователей. Коллаборативная фильтрация представляет собой более универсальный подход и часто дает лучший результат.

Однако для этого вида фильтрации существует проблема холодного старта: о пользователе ничего не известно, когда он впервые входит в систему.

И, наконец, третий вид – гибридная фильтрация – подход, объединяющий два вышеперечисленных подхода.

3.3. Рекомендация конференций

Часто конференции имеют общую тематику «Информационные технологии», и человеку, занимающемуся, например, исследованием в области «Искусственного интеллекта», сложно понять, будет ли эта конференция ему полезна. В связи с этим, формируется следующее требование к системе: система должна уметь автоматически определять тематику конференции.

Действительно, список тем, обсуждаемых на конференции, – одна из самых важных характеристик мероприятия. Так, если на конференции не будет обсуждаться ничего из области научных интересов пользователя, то для него посещение этой конференции может оказаться бесполезным. Поэтому рекомендации в нашей системе строятся в первую очередь по тематике, которая должна соответствовать интересам пользователя. Таким образом, рекомендательный модуль нашей системы является основанным на контенте.

В профиле пользователя имеется информация о его научных интересах, представленная в виде взвешенного дерева терминов, в котором каждой вершине соответствует термин из предметной области и вероятность того, что этот термин отвечает интересам пользователя. Все конференции в системе также представлены в виде аналогичных деревьев. Поэтому мы проводим сравнение дерева пользователя и деревьев конференций и находим наиболее близкие в отношении величины коэффициента сходства. Из этого следует, что перед нами возникает необходимость уметь сравнивать два взвешенных дерева.

3.4. Анализ существующих подходов к сравнению

Существует ряд подходов к вычислению меры сходства двух взвешенных графов. В работе [3] приводится определение нечеткого графа. Под нечетким графом $G = (V, E)$ понимается структура, в которой как вершины $v \in V$, так и ребра $e \epsilon E$ могут быть взвешены либо значением некоторой лингвистической переменной, либо числовым значением от 0 до 1. В нашем случае мы имеем дело с графами, в которых каждой вершине поставлено в соответствие значение из интервала [0; 1]. В задаче определения сходства четких графов ключевую роль играла функция эквивалентности вершин **μ**. Для нечетких графов есть нечеткие аналоги функции импликации и определяющейся через нее функции эквивалентности.

Нечеткая импликация нечетких высказываний α и β – бинарная логическая операция, результат которой является нечетким высказыванием, истинность которого может принимать значение от 0 до 1. Результат истинности этой операции может вычисляться по одной из формул, представленных ниже.

Нечеткая импликация Гёделя:

$T(\alpha \to \beta) = max\ (T(\alpha), T(\beta))$,

где T – функция истинности выражения.

Нечеткая импликация Мамдани:

$T(\alpha \to \beta) = min\ (T(\alpha), T(\beta))$.

Нечеткая импликация Заде:

$T(\alpha \to \beta) = min\ (1 - T(\alpha), T(\beta))$.

Нечеткая импликация Лукасевича:

$T(\alpha \to \beta) = min\ (1, 1 - T(\alpha) + T(\beta))$.

Эквивалентность для четких множеств определялась как

$a_i \leftrightarrow b_i = (a_i \to b_i)\ \&\ (b_i \to a_i)$;

для нечетких множеств аналогом конъюнкции является функция нахождения минимума. Отсюда получаем:

$a_i \leftrightarrow b_i = min\ ((a_i \to b_i), (b_i \to a_i))$.

Сравним значения эквивалентности, полученные с использованием разных формул импликации, на крайних и среднем значениях интервала [0; 1].

Таблица 1.

Эквивалентность вершин a и b при использовании формулы Гёделя

b \ a	0	0,5	1
0	0	0,5	1
0,5	0,5	0,5	1
1	1	1	1

Таблица 2.

Эквивалентность вершин a и b при использовании формулы Мамдани

b \ a	0	0,5	1
0	0	0	0
0,5	0	0,5	0,5
1	0	0,5	1

Таблица 3.

Эквивалентность вершин a и b при использовании формулы Заде

b \ a	0	0,5	1
0	0	0	0
0,5	0	0,5	0
1	0	0	0

Таблица 4.

Эквивалентность вершин a и b при использовании формулы Лукасевича

b \ a	0	0,5	1
0	1	0,5	0
0,5	0,5	1	0,5
1	0	0,5	1

Как видно из таблиц, представленных выше, использование формулы Гёделя приводит к ситуации, когда вершина присутствует только в одном из сравниваемых графов, но тем не менее функция эквивалентности выдает 1. Применительно к нашей системе это означает, что, например, на одной конференции обсуждается тема «Интеллектуальный анализ данных», а на другой конференции эта тема не затрагивается, но в итоге по теме «Интеллектуальный анализ данных» обе конференции признаны эквивалентными. Очевидно, этот вывод противоречит задачам, решаемым нашей системой.

Использование формулы Мамдани приводит к другой проблеме: если вершина отсутствует в обоих графах, то эквивалентность по такой вершине будет нулевой. Однако, исходя из естественных рассуждений, отсутствие вершины в обоих графах говорит об их схожести, а не о различии.

Если использовать формулу Заде, то отсутствие вершины хотя бы в одном из графов приводит к тому, что схожесть таких графов автоматически приравнивается к нулю.

Вышеописанных проблем лишена формула Лукасевича, поэтому было принято решение для расчетов использовать именно её.

Итак, после того, как мы вычислили сходство между всеми соответствующими парами вершин в двух нечетких графах, мы находим собственно меру сходства между этими графами:

$$C(A,B) = \frac{\sum_i (a_i \leftrightarrow b_i)}{n} = \frac{\sum_i (min((a_i \rightarrow b_i),(b_i \rightarrow a_i)))}{n}.$$

3.5. Принцип построения рекомендаций

После того, как пользователь заполнил свой аккаунт и система построила дерево научных интересов пользователя, она начинает построение рекомендательной модели.

Система анализирует все хранящиеся в виде нечетких деревьев конференции и находит меры сходства $C(A, B_i)$ между графом пользователя A и графами конференций B_i.

На следующем этапе система ранжирует все конференции по значению меры сходства и рекомендует пользователю самые подходящие. Таким образом пользователь получает рекомендацию, учитывающую его исследовательские интересы.

4. Вопросно-ответный модуль

В нашей системе молодой исследователь обеспечен возможностью обратной связи. Это означает, что, если у него есть вопросы относительно мероприятий, он может задать их системе и получить ответы на основе информации, которой обладает система.

Эти вопросы могут касаться как организационной информации (время, место, форма участия, стоимость участия), так и тем, обсуждаемых на конференции.

Вопросы понимаются как требование отыскать истинное суждение (или суждение, выполняющееся с удовлетворяющей нас вероятностью). Традиционно в эротетической логике рассматриваются вопросы «какой»-типа и «ли»-вопросы [1].

Однако следует учитывать, что в нашем случае данные имеют нечеткий характер: информация об атрибутах, описывающих конференцию, может отсутствовать в системе, присутствовать, присутствовать с какой-то вероятностью (как в случае с тематикой), или её наличие может быть неизвестно. Необходимо заметить, что традиционная эротетическая логика не учитывает нечеткий характер данных. Поэтому вопросные типы были интерпретированы для нечетких данных, а также были добавлены новые типы [9].

Таким образом, в рамках данной работы была разработана классификация вопросных типов с учетом нечеткого характера данных. Рассмотрим подробнее вопросные типы этой классификации.

«Какой»-вопросы: *Какая форма участия у конференций, проходящих в Новосибирске?*

Такие вопросы нацелены на получение недостающей информации о прецеденте по некоторой известной информации. Другими словами, мы знаем значение одних атрибутов и хотим вычислить значение других с некоторой вероятностью.

Представление этого вопросного типа для вопросов относительно различных атрибутов описано в Табл. 5.

Таблица 5.

Языковые шаблоны для «какой»-вопросов

Название атрибута	Языковой шаблон
Название	Как называется конференция, если верно, что…?
Форма участия	Какая форма участия у конференции, если верно, что…?
Место проведения	Где проводится конференция, если верно, что…?
Время проведения	Когда проводится конференция, если верно, что…?
Стоимость	Какова стоимость участия в конференции, если верно, что…?
Реферативная база данных	Какая реферативная база у конференции, если верно, что…?
Секции	Какие секции представлены на конференции, если верно, что…?
Тематика	Какие темы представлены на конференции, если верно, что…?

«Ли»-вопросы: *Верно ли, что конференции проводятся в Новосибирске и посвящены «Компьютерному зрению»?*

Здесь важно различать вопросы:

1. *Верно ли, что конференции проводятся в Новосибирске и посвящены «Компьютерному зрению»?*
2. *Верно ли, что конференции, которые проводятся в Новосибирске, посвящены «Компьютерному зрению»?*

Второй вопрос – это «ли»-вопрос с надстроенным условием «*которые проводятся в Новосибирске*», то есть для отыскания ответа на вопрос второго типа мы будем рассматривать только ту часть базы, конференции в которой проходили в Новосибирске.

Вероятностные вопросы: Какова вероятность, что конференции проводятся в Новосибирске и посвящены «Компьютерному зрению»?

Аналогично «ли»-вопросам, только в качестве ответа получаем не да/нет, а конкретную вероятностную характеристику выражения.

Условные вопросы: Если известно, что конференция проводится в Новосибирске, то какова вероятность, что конференция посвящена «Компьютерному зрению»?

«Почему»-вопросы являются наиболее трудными для адаптации к данной предметной области, потому что предполагают наличие причинно-следственных связей между атрибутами прецедента-конференции.

Такие связи могут быть неочевидными, а могут и вовсе отсутствовать. Так, например, причина проведения конференции по определенной тематике в Новосибирске может состоять в популярности этой тематики в данном городе. Информация о такой зависимости между атрибутами может быть полезной для организаторов мероприятий, которые находятся в поиске площадок для конференций. Однако знание о зависимости между временем проведения конференции и реферативной базой данных кажется неприменимым.

В связи с этим логично предоставить пользователю следующие возможности:

1. Возможность задавать «почему»-вопросы без всяких ограничений. Ответ будет представлять собой набор значений атрибутов и оценки вероятности того, что эти значения являются причинами:

Почему конференция проводится в Новосибирске? Ответ: секция «Инженерия знаний», вероятность $[\alpha_1; \beta_1]$*; форма участия «очная», вероятность* $[\alpha_2; \beta_2]$*…*

2. В данном случае пользователю предлагается самому оценить полезность таких выводов. Иногда задача состоит именно в том, чтобы найти такие закономерности, поэтому эта интерпретация «почему»-вопросов может оказаться полезной.

3. Возможность задавать «почему»-вопросы с ограничениями. В этом случае пользователь указывает, взаимосвязь с какими атрибутам кажется ему полезной. Такая возможность должна предоставляться интерфейсом.

4. В общем случае вопросы этого типа формируются по правилу:

Почему <тело суждения>?

Заключение

Разработанная система нацелена на то, чтобы помочь молодым исследователям разобраться с огромным разнообразием проходящих конференций по ИТ-тематике и выбрать действительно интересные для них мероприятия. Система анализирует хранящиеся в базе конференции, определяя для каждой из них наиболее вероятные темы обсуждения. Система также анализирует научные интересы пользователя и рекомендует действительно актуальные для него мероприятия. И, наконец, система обеспечивает пользователя возможностью обратной связи.

Таким образом, молодой ученый получает доступ к контексту, в котором он выполняет свою работу: посещая различные конференции, он знакомится с мнением научного сообщества о своем исследовании.

Список источников

1. **Белнап Н., Стил Т.** Логика вопросов и ответов / пер. с англ. Г. Е. Крейдлина, науч. ред. О. Н. Кессиди. М.: Прогресс, 1981. 290 с.
2. **ГОСТ 7.90-2007. Система стандартов по информации, библиотечному и издательскому делу. Универсальная десятичная классификация. Структура, правила введения и индексирования**: издание официальное. М.: Стандартинформ, 2010.

3. **Карелин В. П., Целых А. Н.** Модели принятия решений на основе установления сходства нечетких графов // Известия ЮФУ. Технические науки. 1999. № 3 (13). С. 18-21.
4. **Коршунов А., Гомзин А.** Тематическое моделирование текстов на естественном языке // Труды Института системного программирования РАН. 2012. Т. 23. С. 215-242.
5. **Открытый каталог научных конференций, выставок и семинаров** [Электронный ресурс]. URL: http://konferencii.ru/ (дата обращения: 18.05.2017).
6. **Пальчунов Д. Е., Яхъяева Г. Э.** Нечеткие логики и теория нечетких моделей // Алгебра и логика. 2015. Т. 54. № 1. С. 109-118.
7. **Солтон Д.** Динамические библиотечно-информационные системы. М.: Мир, 1979. 557 с.
8. **Яхъяева Г. Э., Карманова А. А., Ершов А. А., Савин Н. П.** Вопросно-ответная система для управления информационными рисками на основе теоретико-модельной формализации предметных областей // Информационные технологии. 2017. Т. 23. № 2. С. 97-106.
9. **Яхъяева Г. Э., Ясинская О. В., Карманова А. А.** Вероятностная вопросно-ответная система в области компьютерной безопасности // Вестник Новосибирского государственного университета. Серия: Информационные технологии. 2014. Т. 12. Вып. 3. С. 132-145.
10. **Coulter N., French J., Glinert E., Horton T., Mead N., Ralston A., Rada R., Rodkin C., Rous B., Tucker A., Wegner P., Weiss E., Wierzbicki C.** Computing Classification System 1998: Current Status and Future Maintenance: Report of the CCS Update Committee [Электронный ресурс]. URL: http://www.acm.org/about/class/ccsup.pdf (дата обращения: 08.06.2017).
11. **Steyvers M., Smyth P., Rosen-Zvi M., Griffiths T.** Probabilistic Author-Topic Models for Information Discovery // Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Seattle, Washington, 2004. P. 306-315.
12. **The Size of the World Wide Web** [Электронный ресурс]. URL: http://www.worldwidewebsize.com (дата обращения: 18.05.2017).

SYSTEM OF CONFERENCES RECOMMENDATIONS CONSTRUCTION ON THE BASIS OF THE ANALYSIS OF USERS' SCIENTIFIC INTERESTS

Karmanova Anastasiya Aleksandrovna
Tabakov Konstantin Andreevich
Novosibirsk State University
anast.karmy.aa@gmail.com; konkov90@gmail.com

This paper is devoted to description of a conferences analysis system. The developed system automatically determines subject of conferences according to related text documents on the basis of an algorithm developed by the authors. The algorithm is based on the "bag of words" model, the ACM CCS taxonomy and takes into account the hierarchical structure of the dictionary of terms. The developed algorithm is also used when analyzing scientific interests of users. The system constructs individual recommendations. Besides, it contains a question-answer module, which provides users with the opportunity to get feedback.

Key words and phrases: thematic modeling; intellectual analysis of text; recommendatory system; taxonomy; "bag of words"; question-answer systems.

УДК 378.025.7
Педагогические науки

В исследовании ставится вопрос о структуре учебных взаимодействий в образовательном процессе высшей школы. Изучению подвергнут информационный контакт преподавателя со студентами, представленный в терминах «гуманитарное мышление» и «естественнонаучное мышление». В парадигме авторского онтогенетического подхода предложена схема функционального сопряжения двух типов мышления как взаимодействующих противоположностей.

Ключевые слова и фразы: образовательный процесс; знания; понятия; гуманитарное мышление; естественнонаучное мышление; онтогенетический подход; учебные взаимодействия.

Карякин Юрий Васильевич, к.т.н.
г. Томск
art-39-1@yandex.ru

О ДВУХ ТИПАХ МЫШЛЕНИЯ В ОБРАЗОВАТЕЛЬНОМ ПРОЦЕССЕ

Введение. В качестве отправного суждения о явлении «мышление» возьмем следующее: «Мышление – это высшая форма познавательной деятельности человека, социально обусловленный психический процесс опосредованного и обобщенного отражения действительности, процесс поисков и открытия существенно нового» [12]. «Мышление – психический процесс отражения действительности, высшая форма творческой активности человека» [5]. Выберем для дальнейшего анализа родовой признак определяемого понятия – «психический процесс отражения действительности». Этот выбор даёт основание для актуализации мышления