

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
ОДЕССКИЙ НАЦИОНАЛЬНЫЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

КОЛЯДА АНДРЕЙ СЕРГЕЕВИЧ

УДК 004.6:001(043.3/.5)

МОДЕЛИ И МЕТОДЫ ПОИСКА ИНФОРМАЦИИ
В НАУКОМЕТРИЧЕСКИХ БАЗАХ ДАННЫХ

05.13.06 – Информационные технологии

Диссертация на соискание ученой степени
кандидата технических наук

Научный руководитель:
Гогунский Виктор Дмитриевич,
доктор технических наук, профессор

ОДЕССА – 2015

СОДЕРЖАНИЕ

ПЕРЕЧЕНЬ УСЛОВНЫХ СОКРАЩЕНИЙ	4
ВВЕДЕНИЕ	5
РАЗДЕЛ 1	
АНАЛИЗ НАУКОМЕТРИЧЕСКИХ БАЗ ДАННЫХ	10
1.1 Наукометрия и наукометрические показатели	10
1.2 Использование информации из наукометрических баз данных	15
1.3 Характеристика наукометрических баз данных	21
1.4 Структура метаданных научных публикаций и их категоризация	28
1.5 Устройство поисковых систем	34
1.6 Выводы	37
РАЗДЕЛ 2	
РАЗРАБОТКА МОДЕЛИ ИЗВЛЕЧЕНИЯ ИНФОРМАЦИИ ИЗ НАУКОМЕТРИЧЕСКИХ БАЗ ДАННЫХ	39
2.1 Веб интерфейс как основной доступ к информации из наукометрических баз данных	39
2.2 Модель извлечения информации из слабоструктурированных веб страниц	41
2.3 Модель веб скрапинга для автоматизации извлечения данных с веб страниц	48
2.4 Трудности извлечения данных из веб страниц и способы их решения	55
2.5 Гибкое планирование при разработке программного обеспечения	57
2.6 Выводы	64
РАЗДЕЛ 3	
ОБРАБОТКА И АНАЛИЗ ИЗВЛЕЧЕННЫХ ПУБЛИКАЦИЙ	66
3.1 Вероятностное тематическое моделирование	66
3.2 Латентно-семантический анализ извлеченной информации с целью идентификации схожих публикаций	70
3.3 Достоверность идентификации авторства научных публикаций на	

	3
основе латентно-семантического анализа	78
3.4 Модель латентного размещения Дирихле	83
3.5 Выводы	88
РАЗДЕЛ 4	
РАЗРАБОТКА ПРОГРАММНОГО ПРОДУКТА ДЛЯ ИЗВЛЕЧЕНИЯ И ОБРАБОТКИ ИНФОРМАЦИИ ИЗ НАУКОМЕТРИЧЕСКИХ БАЗ ДАННЫХ	90
4.1 Требования и задачи программного продукта	90
4.2 Программный проект	91
4.3 Высокоуровневый дизайн	92
4.4 Детализированный дизайн	98
4.5 Применение и развертывание разработанного программного продукта	112
4.6 Выводы	115
ОБЩИЕ ВЫВОДЫ И РЕКОМЕНДАЦИИ	116
ПЕРЕЧЕНЬ ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	118
ПРИЛОЖЕНИЯ	128

ПЕРЕЧЕНЬ УСЛОВНЫХ СОКРАЩЕНИЙ

БД – база данных

ВТМ – вероятностная тематическая модель

НБД – наукометрическая база данных

ПО – программное обеспечение

API – интерфейс программирования приложений (Application Programming Interface)

LDA – латентное размещение Дирихле (Latent Dirichlet Allocation)

LSA – латентно-семантический анализ (Latent Semantic Analysis)

SVD – сингулярное разложение (Singular Value Decomposition)

ВВЕДЕНИЕ

Актуальность диссертационного исследования. Развитие информационных технологий по организации международных наукометрических баз данных и электронных библиотек порождает новые возможности и задачи в сфере образовательной и научной деятельности в высшей школе Украины. Одним из направлений этой деятельности является определение обобщенной оценки качества и результатов научных исследований отдельного ученого, кафедры, университета и высших учебных заведений Украины в целом [1 – 2]. Анализ характеристик и основных свойств НБД и индикаторов цитирования научных публикаций отражают широкий спектр целевого назначения НБД – от обеспечения сугубо информационных потребностей ученых до всестороннего анализа публикационной активности авторов. Определение наиболее применяемых характеристик производительности научной деятельности в мировом научном сообществе для ознакомления широкого круга ученых с показателями оценки значимости научных публикаций является актуальной научно-прикладной проблемой.

На сегодняшний день Министерство образования и науки Украины целенаправленно ориентирует публикационную деятельность ученых на вхождение в мировое научное сообщество. Публикационная активность научных сотрудников является одним из основных факторов, который учитывается при определении мировых рейтингов высших учебных заведений. При этом НБД являются основными каналами дальнейшего применения научных результатов, как главной информационной и социальной характеристики страны, университета, научного коллектива или отдельного ученого. Поэтому работа по извлечению информации из них является достаточно актуальной проблемой.

Существующие НБД, как правило, ориентированы на поиск публикаций только в своих хранилищах. При этом различные НБД используют свои

специфические формы интерфейса, что определяет необходимость обязательной личностного участия ученых в поиске публикаций в различных базах. Эти обстоятельства порождают противоречие между необходимостью интегральной оценки публикационной активности авторов и отсутствием информационных технологий, позволяющих выполнять информационно-поисковые операции в различных НБД. Кроме того существует задача многовариантного задания атрибутов поиска, в том числе, различных вариантов написания фамилий, поскольку некоторые НБД используют только английские варианты фамилий. В разных НБД не используются модели, методы и инструментальные методы определения достоверности полученной информации, основанные на анализе скрытых переменных для выявления связей в наборе названий публикаций, позволяющей достоверно идентифицировать публикации конкретных авторов. Поэтому разработка программных продуктов для извлечения информации из различных НБД является актуальной задачей, которая направлена на решение указанных противоречий в области информационных технологий.

Связь работы с научными программами, планами, темами.

Диссертационная работа выполнялась в соответствии с планами научных исследований ОНПУ, с участием автора в качестве исполнителя, по планам «Теория и практика компетентностного управления персоналом в организационно-технических и социальных системах» (ДР № 0113U007624, 2012-2015) и «Методология внедрения проектно-векторного управления информационными средами для мониторинга и управления научными исследованиями научных групп» (ДР № 0115U000330, 2015).

Цель и задачи исследования. Целью исследования является разработка моделей, методов и информационной технологии поиска информации для повышения точности поиска публикаций в наукометрических базах данных. Для достижения поставленной цели в диссертации поставлены и решены следующие задачи:

- анализ структуры метаданных и способов их представления в наиболее распространенных и известных наукометрических базах данных;
- анализ структуры поисковых систем с акцентом на инструменты загрузки веб-страниц, и преобразования найденной информации в структурированный формат данных;
- разработка метода получения информации из самых распространенных наукометрических баз данных с возможностью расширения их количества;
- разработка модели хранения метаданных для дальнейшей их обработки;
- разработка методов ранжирования информации с целью поиска наиболее релевантных результатов;
- разработка информационно-поисковой системы с удобными интерфейсами для задания критериев поиска и получения результатов.

Объект исследования – процесс поиска информации в наукометрических базах данных.

Предмет исследования – модели, методы и инструментальные средства для создания информационно-поисковых систем.

Методы исследования. Для достижения цели и решения задач, поставленных в диссертационной работе, используется теоретический анализ способов доступа к информации из наукометрических баз. Метод имитационного моделирования используется при построении вероятностных моделей со скрытыми переменными (для идентификации публикаций конкретного автора).

Научная новизна полученных результатов.

- впервые разработан комплексный метод поиска и преобразования информации с наукометрических баз данных в структурированный формат данных, что позволяет уменьшить время на интеграцию новых баз в систему;
- впервые разработаны методы ранжирования публикаций по их названию, с использованием вероятностных моделей латентного семантического анализа и размещения Дирихле, что позволяет повысить точность поиска;

– впервые разработана модель информационно-поисковой системы без предварительного сканирования полного содержания наукометрических базы данных, что позволяет сократить использование вычислительных ресурсов;

– получили дальнейшее развитие инструментальные средства информационных технологий поиска и обработки информации с динамически созданных веб страниц на стороне пользователя (клиента), что значительно расширяет число пригодных к обработке документов.

Практическая значимость полученных результатов. Разработан программный продукт «Science Metric Databases» (SMD), реализующей информационную технологию поиска данных в различных наукометрических базах данных. SMD, как составляющая информационно-поисковой системы метаданных публикаций ученых вузов Украины, использованная в ГДР «Методология внедрения проектно-векторного управления информационными средами для мониторинга и управления научными исследованиями научных групп». Система позволяет повысить точность поиска публикаций отдельного автора по сравнению с похожими системами или поисковыми машинами. Открытость архитектуры позволяет расширять возможности системы за счет интеграции других наукометрических баз данных. Полученные метаданные публикаций используются программным комплексом для мониторинга групповой публикационной активности лабораторий, кафедр и институтов. Система развернута на сервере кафедры Управления системами безопасности жизнедеятельности ОНПУ.

Личный вклад соискателя. Все результаты научных, теоретических и практических исследований, изложенные в диссертации, получены автором самостоятельно. В публикациях, опубликованных в соавторстве, использовались только те положения и идеи, которые являются результатом личных достижений аспиранта. Автором разработана информационно-поисковая система, которая является результатом выполнения данной работы.

Апробация результатов диссертации. Основные результаты работы докладывались и обсуждались на конференциях и семинарах: международная научно-практическая конференция (МНПК) «Управление проектами: состояние и перспективы» (Николаев 2013); научно-методический семинар «Пути реализации кредитно-модульной системы» (Одесса 2011, 2013, 2014, 2015); научный семинар «Информационные технологии в образовании, науке и производстве» (Херсон 2012, 2014); МНПК «Автоматизация: проблемы, идеи, решения» (Севастополь 2014), МНПК «Молодежь в мире современных технологий по тематике: теоретико-методологические и научно-практические основы управления проектами повышения конкурентоспособности территорий» (Херсон 2014).

Публикации. Основные положения и результаты диссертационной работы отражены в 18 публикациях, из них 7 статей в научных изданиях Украины, которые также индексируются в наукометрических базах данных, 8 статей в научно-технических сборниках и 3 тезиса докладов конференций.

Структура диссертации. Диссертация состоит из введения, четырех глав, заключения, списка использованных источников из 83 наименований на 10 страницах, 2 приложений. Общий объем основной части составляет 113 страниц текста, в том числе содержит 33 рисунка (один из которых занимает всю страницу) и 13 таблиц (3 из которых занимают всю страницу).

РАЗДЕЛ 1

ПРОБЛЕМЫ АНАЛИЗА НАУКОМЕТРИЧЕСКИХ БАЗ ДАННЫХ

1.1 Наукометрия и наукометрические показатели

Теоретические, функциональные и структурные изменения в различных областях знаний определенным образом отражаются в научных публикациях. Именно множество публикаций является основой для формирования новых знаний. Мирской опыт взаимодействия сообщества ученых в информационной среде всемирной Web-паутины свидетельствует о целесообразности применения некоторых показателей производительности научной деятельности [19–21]. Поэтому актуальной задачей является разработка информационных технологий для учета публикаций ученых Украины в зарубежных журналах или в отечественных изданиях, которые индексируются в международных наукометрических базах [22].

Идеология современной наукометрии базируется в основном на работах Derek J. deSollaPrice и Eugene Garfield [23–24]. Последний основал Институт научной информации (приобретенный Thomson Scientific в 1992 году), который активно проводит исследования в области наукометрического анализа. Появление наукометрии стало следствием экспоненциального роста науки в середине XX века, когда было обращено внимание на изменение характера научных исследований – научные сотрудники вынуждены тратить почти 50% своего рабочего времени на информационную деятельность.

Задачи наукометрии решаются специализированными институтами и службами. Из множества изученных и испытанных наукометрических методов для решения информационных задач пользователя наиболее подходят следующие методы:

– статистический – использует такие показатели, как количество ученых, журналов, заказов на годовые комплекты журналов в библиотеках и информационных центрах;

– подсчет числа публикаций – измеряется количество научных продуктов (книг, статей, отчетов и др.), объединяемых общим термином – «публикация». Это позволяет судить об актуальности и перспективности данного научного направления, и способствуют рациональному предпроектному оппонированию научной деятельности именно на стадии планирования своей работы, выбора актуальной темы своего будущего исследования. Распределение ученых по количеству публикаций позволяет не только выявить производительность, но и определить ранг ученого, и, следовательно, его значимость;

– индекс цитирования – базируется на обязательности ссылок в научных публикациях; в его основе лежит наукометрический индикатор – число ссылок. Этот метод используется для измерения научной значимости результатов деятельности ученого - научных публикаций.

В сфере наукометрических измерений используются специализированные информационные объекты, которые называют наукометрическими базами данных. Они представляют собой средства хранения и обработки наукометрических показателей, а также, зачастую, содержат тело публикационных материалов (статей, журналов, книг), то есть являются научным репозиторием. Наиболее авторитетные наукометрические базы данных мирового уровня (Scopus, Web of Science) являются реферативными базами данных, которые включают в себя некоторое конечное множество публикаций, а также средства сервиса для удовлетворения информационных потребностей конечных пользователей. При этом, за некоторым исключением, предоставление информационных услуг осуществляется на платной основе.

Наукометрические показатели разделяют на следующие категории [25]:

– на основе количества публикаций (суммарное число, взвешенная сумма, отношение количества публикаций к научному стажу);

– на основе количества цитирования (суммарное число ссылок, количество скрытого цитирования);

– на основе количества публикаций и количества цитирования (индекс Хирша и различные его модификации).

Следует отметить, что эти показатели определяются с некоторой погрешностью, это связано с ошибками в пристатейных списках литературы, сокращением списка первоисточников, применением различных стандартов описания использованных источников, забывание имен классиков и др.

Количественные оценки основаны на опубликованных данных – это число публикаций, анализ частоты их цитируемости (индексы цитирования), h-индекс или индекс Хирша, импакт-фактор научного журнала, в котором работы опубликованы. Из перечисленных показателей в последнее время чаще всего применяются: индекс цитирования, h-индекс и импакт-фактор.

Импакт-фактор (ИФ или IF) – численный показатель научного уровня журнала. С 1960-х годов он ежегодно рассчитывается Институтом научной информации (англ. Institute for Scientific Information, ISI) и публикуется в журнале «Journal Citation Report». Согласно ИФ оценивают уровень журналов, качество статей, опубликованных в них, дают финансовую поддержку исследователям и принимают сотрудников на работу. Импакт-фактор имеет хотя и большое, но неоднозначно оцениваемое влияние на оценку результатов научных исследований.

Расчет импакт-фактора основан на трехлетнем периоде. Например, импакт-фактор журнала в 2013 году I_{2013} вычисляется следующим образом:

$$I_{2013} = \frac{A}{B}, \quad (1.1)$$

где A – число цитирований в течение 2013 года в журналах, отслеживаемых Институтом научной информации, статей, опубликованных в данном журнале в 2011-2012 годах;

B – число статей, опубликованных в данном журнале в 2011-2012 годах.

Индекс Хирша (h-индекс) – наукометрический показатель, предложенный американским физиком Хорхе Гиршем [26]. Индекс Хирша может применяться для количественной характеристики производительности одного ученого, группы ученых, кафедры, факультета, университета или страны в целом, что определяется на основе количества статей и числа цитирований этих публикаций. Для определения индекса Хирша имеющиеся статьи ранжируют по нисходящему числу ссылок на них. Далее, определяют ту статью, ранг которой совпадает с числом ее цитирований. Это число и является h-индекс, который определяется точкой пересечения кривой ранжированных цитирований и линией $z = r$ (45 grad).

Индекс Хирша может исчисляться с использованием как общедоступных наукометрических баз данных в Интернете (например, Google Scholar, Science Index), так и баз данных с платной подпиской (например, Scopus или Web of Science). Следует отметить, что индекс Хирша, подсчитанный для одного и того же ученого с использованием различных баз данных, будет, вообще говоря, разным – как и другие наукометрические характеристики. Он зависит от области охвата выбранной базы данных, как по объему статей в БД, так и интервалов времени по которому учитываются статьи.

Индекс Хирша может определяться с учетом и без учета самоцитирования. Считается, что отбрасывание ссылок авторов на собственные статьи дает более объективные результаты.

Под индексом цитирования понимается реферативная база данных научных публикаций, которая индексирует ссылки, указанные в пристатейных списках этих публикаций и предоставляющая количественные показатели этих ссылок (такие как суммарный объем цитирования, индекс Гирша и др). Из статей в журналах, включенных в реферативную базу извлекаются традиционная библиографическая информация (исходные данные) и списки цитируемой литературы (пристатейная библиография) [2].

Доступ к множеству публикаций мирового сообщества ученых формирует новое отношение к такой слабо структурированной области как базы данных научных публикаций. Именно публикации является отражением результатов научных достижений, а цитирование этих публикаций коллегами является оценкой качества и значимости для науки и практики полученных результатов.

Увеличение показателей цитирования научных публикаций украинских ученых в наукометрических базах данных связано с активной деятельностью в таких направлениях: расширение ареала публикаций в уже существующих научных изданиях, включенных в наукометрические базы; включения большего количества научных изданий Украины в наукометрические базы.

При этом особое внимание следует уделить качеству научных публикаций – не только с точки зрения новизны и практической значимости исследований, но и в плане представления текста статей приемлемой научной английской языке. Научные издания, не соответствуют требованиям, следует переформатировать и привести в соответствие с принятыми правилами представления научных публикаций в иностранных репозиториях и международных наукометрических базах данных.

С развитием информационных технологий наукометрические базы данных получили автоматизированные средства обеспечения их работы, повышение удобства пользования, расширение функционала и широкое распространение.

До недавнего времени о наукометрических базах знали и использовали их в основном библиотечные работники различных стран, в том числе и Украины. Сегодня же, как минимум, каждый научный сотрудник Украины знает о них, так как существуют требования на публикацию статей в изданиях, которые включены в международные наукометрические базы данных. Это также поможет заменить перечень профильных изданий ВАК Украины, который периодически обновлялся в ручном режиме, а значит, требовал постоянного внимания официальных инстанций, на более прозрачный по принципу

формирования, автоматически обновляющийся список журналов, включенных в международные наукометрические базы [55].

Национальная библиотека Украины имени В. И. Вернадского в качестве примера того, какую информацию можно извлечь из международных наукометрических баз данных приводит различные рейтинги, составленные по показателям наукометрических баз данных. Также они создали отдел наукометрии и библиометрики, репозиторий которого не имеет аналогов в Европе и СНГ среди открытых проектов. [25]

«Пан Бібліотекар» приготовил специально для украинских ученых актуальные на сегодня списки украинских журналов в международных наукометрических базах (Scopus, Web of Science) [56, 57]. Также составлен список основных баз данных и поисковых систем, которые могут быть полезны в академической среде для поиска и доступа к статьям в научных журналах, либо в хранилищах, архивах и других сборниках научных статей

1.2 Использование информации из наукометрических баз данных

Публикационная активность ученых является одним из основных факторов, который учитывается при определении мировых рейтингов высших учебных заведений. С личной заинтересованности ученых публикации превращается в реальный весомый показатель деятельности вуза.

Влиятельный рейтинг лучших университетов мира (QS World University Rankings) определяется по методике британской консалтинговой компании Quacquarelli Symonds (QS) [20]. Он считается одним из самых влиятельных глобальных рейтингов университетов. Предложенный в 2004 году Quacquarelli Symonds совместно с британским изданием Times Higher Education. К 2010 году был известен как The World University Rankings. С 2010 года единый рейтинг разделился на два: издание Times Higher Education выпускает рейтинг лучших университетов мира The World Reputation Rankings совместно с агентством

Thomson Reuters, a Quacquarelli Symonds продолжает выпускать рейтинг под названием QS World University Rankings.

Рейтинг оценивает университеты по следующим показателям: активность и качество научно-исследовательской деятельности, выводы работодателей и карьерный потенциал, качество обучения и интернационализация. Эти показатели охватывают ключевые стратегические миссии университетов мирового значения, за которые они отвечают перед участниками процесса: академическим сообществом, работодателями, учащимися и их родителями. Ежегодно в исследовании оцениваются более 2,5 тысячи высших учебных заведений по всему миру. По его итогам составляется рейтинг 500 лучших университетов мира, а также рейтинги университетов по отдельным дисциплинам.

Уровень достижений университетов оценивается на основании результатов комбинации статистического анализа деятельности учебных заведений, аудированных данных (включая информацию об индексах цитирования из базы данных Scopus, крупнейшей в мире библиометрические базы данных научных публикаций), а также данных глобального экспертного опроса представителей международного академического сообщества и работодателей, которые выражают свои мысли об университетах.

Опросу привлекаются представители международного академического сообщества (QS Global Academic Survey), преимущественно профессора и руководители университетов. Среди них ведущие ученые и ректоры более 500 университетов. Респондент может назвать до 30 университетов, не называя вуза, в котором он работает. Опрос проводится в пяти предметных областях: естественные науки, социальные науки, гуманитарные науки и искусство, науки о жизни, инженерные науки и технологии. В опросе представителей работодателей (QS Global Employer Survey) участвуют несколько тысяч компаний из более чем 90 стран мира. Ниже представлены показатели, по которым проводится оценка деятельности университетов (рис. 1.1).

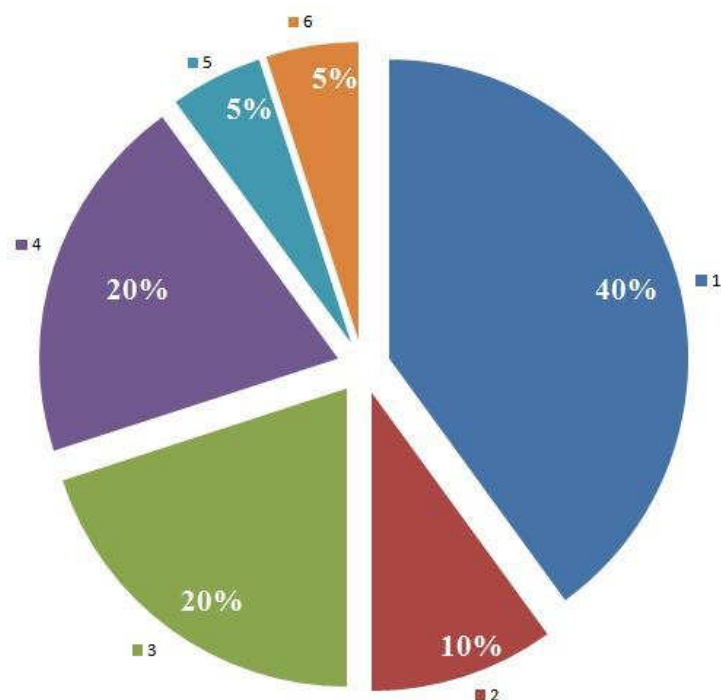


Рисунок 1.1 – Соотношение составляющих QS рейтинга:

- 1 – индекс академической репутации (опрос);
- 2 – индекс репутации среди работодателей (опрос);
- 3 – соотношение профессорско-преподавательского состава по отношению к численности учащихся;
- 4 – индекс цитирования научных статей преподавательского состава по отношению к численности преподавательского состава (Scopus);
- 5 – доля иностранных преподавателей по отношению к численности преподавательского состава (по эквиваленту полной ставки);
- 6 – доля иностранных студентов относительно численности учеников (программы полного цикла обучения).

Другой влиятельный рейтинг лучших вузов мира ARWU (Академический рейтинг университетов мира) состоит Институтом высшего образования Шанхайского университета Цзяо Тун [21]. Поэтому его часто называют Шанхайским рейтингом. Результаты рейтинга в виде списка 500 лучших вузов мира публикуются с 2003 года. Ниже представлены показатели оценки деятельности университетов по Шанхайским рейтингом (рис. 1.2).

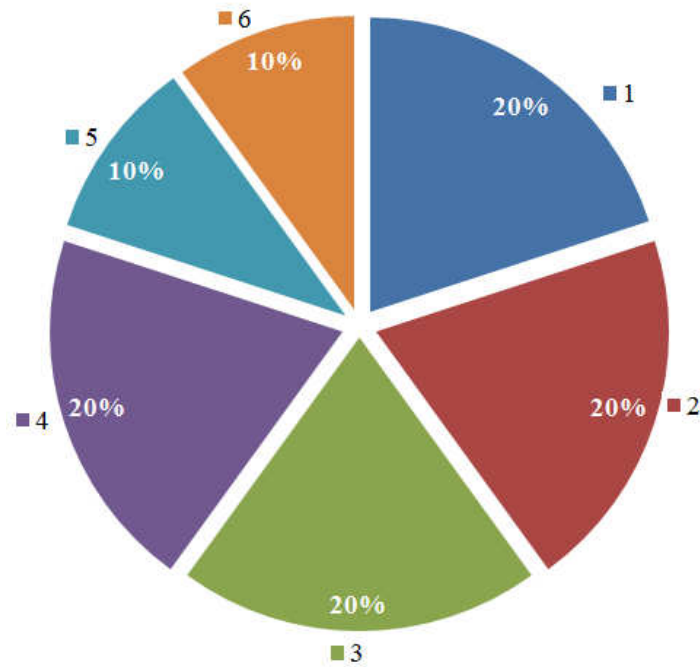


Рисунок 1.2 – Показатели по Шанхайским рейтингам:

- 1 – количество статей, опубликованных в Nature или Science;
- 2 – число часто цитируемых публикаций (показатель цитируемости SCIE - Science Citation Index – Expanded и SSCI – Social Science Citation Index);
- 3 – число преподавателей, получивших Нобелевскую или Филдсовскую премии;
- 4 – число часто цитируемых в научных изданиях публикаций;
- 5 – число выпускников ВУЗа, получивших Нобелевскую или Филдсовскую премию;
- 6 – соотношение пяти вышеизложенных показателей и численности персонала ВУЗа.

При составлении Шанхайского рейтинга лучших вузов мира отбираются только те университеты, преподаватели или выпускники которых имеют Нобелевскую или Филдсовскую премию, публикуют в научных изданиях цитируемые научные исследования, индексируемые в Science Citation Index Expanded и Social Science Citation Index. Таких вузов получается всего 1000 в мире, 500 лучших из них попадают в Шанхайского рейтинга.

К сожалению, украинские университеты в этих рейтингах значительно отстают от ведущих университетов США, Канады, Англии и Германии.

Тенденции экономики современного информационного общества таковы, что движущей силой инновационного развития общества становится наука [?].

Можно отметить, однако, что МОН Украины целенаправленно ориентирует публикационную деятельность ученых на вхождение в мировое научное сообщество. Распространение и продвижение научных достижений осуществляется через информационные каналы, среди которых, как наиболее значимые, можно выделить средства массовой информации, Интернет, мобильные технологии. Наукометрические базы данных являются основными ячейками трансформации знаний и каналами дальнейшего применения научных результатов, как главной информационной и социальной характеристики страны, университета, научного коллектива или отдельного ученого.

Сегодня уровень наукоемкости и совершенства систем различного назначения определены в мире как ключевой механизм формирования конкурентоспособности государства и бизнеса. Поэтому насущной задачей для Украины является побуждение ученых к публикации результатов своих исследований в зарубежных журналах, или в изданиях Украины, включенных в зарубежных наукометрических баз. Можно выделить следующие мероприятия в этом направлении:

- изменены требования к научным изданиям для включения в перечень профессиональных изданий с целью создания условий соответствия этих изданий международным требованиям. В связи с новыми требованиями к научным публикациям, а также на основе анализа нынешнего состояния научных изданий можно сделать вывод, что очень маленькая часть научных изданий соответствуют требованиям, которые существуют в международных наукометрических базах данных;

- конкурс по финансированию проектов научных исследований и разработок с 2013 проводится с учетом числа публикаций в Scopus и других международных наукометрических базах;

- ужесточены требования «к планированию диссертационных

исследований, формулировка их тематики, в частности по предотвращению высказываний общего при формулировании темы, новизны, предмета и объекта исследования»;

– существенной составляющей диссертационных исследований становятся публикации полученных результатов в специализированных профессиональных изданиях, а также в электронных изданиях. Новые требования к публикации результатов диссертации следующие: докторская диссертация – всего не менее 20 профессиональных публикаций, из них «не менее четырех публикаций в научных периодических изданиях других государств по направлению, с которого подготовлен диссертацию»; кандидатская диссертация – всего не менее 5 профессиональных публикаций и «не менее одной статьи в научных периодических изданиях других государств по направлению, с которого подготовлен диссертацию». К публикациям в научных периодических изданиях других государств могут приравниваться публикации в профессиональных изданиях Украины, включенных в международных наукометрические базы».

– к критериям оценки деятельности высших учебных заведений включен показатель «численность научно-педагогических работников, имеющих публикации в изданиях иностранных государств или в изданиях Украины, включенных в международных научно метрических баз в отчетном учебном году».

Указанные требования к научным публикациям по результатам диссертационных исследований, непосредственное оценивание ВУЗов по количеству публикаций в изданиях иностранных государств и научных изданиях Украины, входящих в международных наукометрические базы, а также формирование новых государственных требований по аккредитации, трансформируют публикационную активность ученых высших учебных заведений с личной заинтересованности профессорско-преподавательского состава в один из важнейших показателей деятельности вуза. Это означает, что планирование набора абитуриентов, выборы преподавателей, научных

исследований будут основываться на данных о публикациях и показатели цитирования. Поэтому, очевидно, для управления этим процессом следует создать инструменты «измерения» публикационной активности ученых с последующим формированием информационно-аналитической системы мониторинга публикационной активности ученых высших учебных заведений Украины [1-2, 32].

Авторы [10] отмечают, что «не снимая с аспирантов, соискателей ученых степеней и всех ученых вузов задач о необходимости представления статей в зарубежные издания следует все же отметить, что и перед высшим менеджментом университетов и МОН встает проблема создания благоприятных условий для включения большего числа изданий Украины с международными наукометрические базы». И хотя этот процесс довольно длительный и затратный, надо неустанно двигаться в этом направлении.

1.3 Характеристика наукометрических баз данных

Международная практика наукометрических исследований сегодня наиболее часто базируется на использовании двух баз данных: Web of Science и Scopus. Широко известны также другие базы данных, которые ориентированы на информационное обеспечение научных исследований без формирования данных наукометрии (табл. 1.1). Все они в основном не являются коммерческими базами.

Среди некоммерческих наукометрических баз, в которых индексируются публикации украинских ученых, можно назвать следующие [2]:

Copernicus – www.journals.indexcopernicus.com/search_article.php;

BASE – www.base-search.net;

DOAJ – <http://www.doaj.org>;

Driver – www.driver-repository.eu;

FreeFullPDF – www.freefullpdf.com;






Science Index – <http://elibrary.ru>;









UlrichsWeb – www.ulrichsweb.com/ulrichsweb/faqs.asp и другие.

Краткая характеристика широко известных международных наукометрических баз данных приведена в табл. 1.1, где рассмотрены объем данных БД, отрасль исследований и наличие наукометрических данных.

В Украине часто используют наукометрическую базу Science Index проекта российской электронной библиотеки eLibrary.ru (РИНЦ - Российский Индекс Научного Цитирования), которая включает 46470 научных журналов, из них только 8621 российских изданий, а остальные – журналы других стран.

Таблица 1.1 – Характеристика международных наукометрических баз

№	Название НБД	Объем данных	Отрасль данных	Наукометрия
1	 Издательство Elsevier, www.scopus.com доступ платный	50 млн. записей из 21 тис. зданий, 370 серий книг, 5,5 млн. тезисов конференций, 25,2 млн. патентов	Наука про жизнь; здоровье; физические науки; социология и гуманитарные (42 издания Украины)	Полные тексты, комплекс данных наукометрии
2	Science Direct Издательство Elsevier, www.sciencedirect.com/ доступ свободный	>12 млн. статей, >3300 журналов и почти 20,000 книг из НБД Scopus	Наука про жизнь; здоровье; физические науки; социология и гуманитарные	Полные тексты, комплекс данных наукометрии
3	Web of Science thomsonreuters.com/web-of-knowledge/ доступ платный	> 12,000 журналов, В том числе в свободного доступа	Многопрофильная база данных	Полные тексты, комплекс данных наукометрии
4	 Издательство Springer link.springer.com доступ платный	>8 млн. журналов и книг	Многопрофильная база данных	Полные тексты, рефераты
5	 РОССИЙСКИЙ ИНДЕКС НАУЧНОГО ЦИТИРОВАНИЯ elibrary.ru/ доступ свободный	46470 журналов, (8621 журнала из РФ); >18,6 млн. статей, 6500 журнала из Scopus	Многопрофильная база данных РИНЦ, 546 журналов Украины	Полные тексты, комплекс данных наукометрии
6	 www.copernicus.org/ доступ свободный	>5000 изданий, в том числе, >1200 журналов в Польше	Многопрофильная база данных, 95 журналов Украины	Рефераты, URL полных текстов, импакт фактор изданий и статей
7	 Bielefeld Academic Search Engine www.base-search.net/ доступ свободный	>52,5 млн. статей з 2776 источников	Многопрофильная база данных, 36 журналов Украины	Рефераты, URL полных текстов

8	 www.doaj.org/ доступ свободный	>1,5 млн. статей 9979 журналов 147 стран	Многопрофильная база данных, 18 журналов Украины	Рефераты, пересылка на полные тексты
9	 University of Michigan Library http://www.lib.umich.edu/ доступ свободный	>500 млн статей, >134 тис видань (у тому числі WoS)	Многопрофильная база данных	Рефераты, пересылка на полные тексты
10	 Дублин, штат Огайо, США www.worldcat.org/search	>240 млн. записей на 470 языках, 72 тис. библиотек з 170 стран.	Многопрофильная база данных, 6 журналов Украины	Библиографическая база, пересылка на полные тексты
11	 Sophia Antipolis technology park www.freefullpdf.com/ доступ свободный	>80 млн. свободных научных публикаций	Многопрофильная база данных	Библиографическая база, пересылка на полные тексты
12	AGRIS agris.fao.org доступ свободный	>5 млн. записей, >22 тис журналов >200 библиотек з >100 стран	Agricultural Sciences and Technology	Библиографическая база, пересылка на полные тексты
13	 доступ свободный	Open access to 918710e-prints	Многопрофильная база данных	Полные тексты
14	 http://www.ans.org/pi/ платный доступ	>500 тис записей	Атомная энергетика	Полные тексты
15	GetInfo www.tib-hannover.de/ доступ свободный	>160 млн. статей	Многопрофильная база данных	Пересылка на полные тексты
16	 worldwidescience.org/ доступ свободный	>100 млн. записей 100 баз данных	Многопрофильная база данных научных публикаций	Полные тексты, пересылка на полные тексты
17	 scimagojr.com/	НБД БД по данным Scopus	Многопрофильная база данных	Комплекс данных наукометрии
18	Registry of Open Access Repositories (ROAR) roar.eprints.org/	Больше 230 стран мира	Сайт показателей цитирования научных изданий	Пересылка на полные тексты
19	Google Академія scholar.google.com.ua/ доступ свободный	Все открытые источники Internet: библиотеки, репозитории	Поисковая и наукометрическая база данных	Пересылка на полные тексты, индекс Хирша
20	UlrichsWEB ulrichsweb.serialssolutions.com доступ свободный	Серийные издания мира (ISSN)	БД регистрационных данных серийных изданий мира	URL серийных изданий с полными текстами

Science Index позиционируется как информационно-аналитическая система, аккумулирует больше 18600000 публикаций, а также информацию о цитировании этих публикаций. С 2014 в Science Index доступна часть изданий, индексируемые в Scopus, 6500 журналов. В проекте Science Index (eLibrary.ru) все публикации в течение года после издания можно получить по заказу за определенную плату. Через год, если это разрешено издателем, все статьи находятся в свободном доступе.

Заслуживают внимания также поисковые системы: Google Академия (<http://scholar.google.com.ua/>) и Publish or Perish [29].

Google Академия является поисковой системой и вместе с тем открытой наукометрической базой данных научных публикаций одновременно. Google Академия является составной частью браузера GoogleChrome и поддерживается компанией Google. В определенной степени она является некоммерческим аналогом Scopus, но с меньшим набором параметров наукометрии. Она позволяет выполнять расширенный поиск публикаций (по фамилии автора или по названию статьи) в научных источниках, которые есть в свободном доступе в Интернет-пространстве. Кроме того, она определяет частоту цитирования найденных по запросу публикаций. После простой регистрации ваш онлайн источник индексируется. Для проверки достаточно ввести его название или фамилию автора в строку запроса GoogleScholar, чтобы получить результаты поиска. Система вычисляет индекс Хирша публикациями автора, а также отражает количество цитирований каждой статьи. Google Академия предоставляет возможность всем авторам научных публикаций создать частную Web-страницу, в которой аккумулируются все статьи автора и отражаются наукометрические данные. Простота регистрации и интуитивно понятный интерфейс определили широкое использование этой системы среди учёных. Сейчас Google Академия предлагает ученым новые информационные услуги: "Моя библиотека" и "Проверено мной". На Web-странице "Моя библиотека" аккумулируются данные о научных статьях и книги с той научной

отрасли, в которой работает автор. Web-страница "Проверено мной" является полезной при написании статей, а также при анализе направлений исследований других авторов, которые также цитируют те же статьи.

Publish or Perish также является бесплатной научной поисковой системой, которая разработана профессором в области международного менеджмента Anne-Wil Harzing (Австралия) при поддержке компании Google. Программу Publish or Perish надо скачать из Интернета и установить ее. Эта поисковая система позволяет выполнять поиск публикаций по фамилии автора. Кроме того, можно определить импакт-индексы цитирования для научных журналов – по числу цитирований статей определенного научного издания исчисляется индекс Хирша этого издания. Результатом работы системы является полный комплект наукометрических показателей по публикациям автора – от индекса Хирша к числу соавторов в найденных статьях.

Как уже упоминалось, наиболее известные и крупные наукометрические базы это Scopus и Web of Science [56, 57].

Scopus – крупнейшая реферативная база данных рецензируемой литературы, включающая инструменты для отслеживания, анализа и визуализации исследований. Scopus обеспечивает наиболее полный обзор мирового производства научных работ в области науки, техники, медицины, социальной науки и искусств и гуманитарных наук.

Scopus обновляется ежедневно и включает в себя:

- 21000 наименований от более чем 5000 международных издательств:
- 20000 рецензируемых журналов (в том числе 2600 журналов открытого доступа)
- 390 отраслевые издания
- 370 серий книг
- 5.5 млн. докладов с конференций
- "Статьи-в-Прессе" из более чем 3850 журналов и издательств таких как Cambridge University Press, Elsevier, Springer, Wiley-Blackwell, издательская

группа Природа и Институт инженеров по электротехнике и радиоэлектронике (ИЕЕ).

50 млн. записей Scopus состоят из:

- 29 млн. записей, в том числе ссылок, до 1995 года (84% включают в себя тезисы)
- 21 млн. записей до 1996, датирующихся начиная с 1823года

Также Scopus индексирует 25.2 млн. патентов из пяти патентных офисов (США, Европы, Японии, Великобритании).

Scopus также предоставляет широкие возможности поиска (по документу, автору, учреждению и т.д.) и анализа (цитирование с привязкой ко времени, индекс Хирша, различные метрики журналов).

Web of Science – наиболее авторитетный индекс цитирования во всем мире, покрывающий ведущую научную литературу. Web of Science доступен через Web of Knowledge – платформу всесторонних исследований, которая используется более 5600 учреждениями в более чем 100 странах.

Web of Science предоставляет:

- Охват более 12000 журналов;
- Более 148 000 трудов наиболее значимых конференций, симпозиумов, семинаров, коллоквиумов и конвенций по всему миру;
- Охват литературы с 1900 года;
- Прямые ссылки на полный текст и информацию библиотечных фондов;
- Инструменты идентификации автора;
- Возможности анализа и инструменты визуализации.

Используя поиск в Web of Science можно найти истинный индекс цитирования по всем дисциплинам и отрезкам времени по интересующей теме (влияние вашей работы, ключевые исследования в данном направлении). А с помощью средств анализа можно проанализировать результаты по автору, дате публикации, учреждению, языку и получить результаты в численном, процентом или графическом представлении.

В последнее время набирает популярность академическая социальная сеть Mendeley [64]. Это также и бесплатная программа для управления библиографической информацией, позволяющая хранить и просматривать исследовательские труды в формате PDF, а также имеющая подключение к международной социальной сети учёных. Для получения доступа к использованию программы, необходимо создать учётную запись на сайте социальной сети. Базовый пакет Mendeley распространяется бесплатно, однако существуют платные версии с увеличенными квотами на хранение материалов и создание групп.

Возможности программы:

- автоматизированное извлечение метаданных из документов PDF;
- синхронизация с учётной записью;
- встроенный просмотрщик PDF с возможностью текстовых пометок (аннотаций).
- поиск по всей библиотеке (по названию документа, имени автора или собственным ключевым словам).
- автоматическое управление PDF-файлами (переименование согласно заданной схеме, мониторинг папок — англ. watch folder).
- поиск недостающей метаинформации через Google Scholar.
- экспорт частей библиотеки в формате BibTeX.
- извлечение сносок из раздела ссылок («References»).
- использование тегов для категоризации документов.

Возможности социальной сети Mendeley:

- статистика просмотра документов.
- 2 ГБ серверного пространства для хранения документов.
- создание профиля с указанием интересов и прочей личной информации.
- букмарклет для автоматического импортирования документа в библиотеку из сайтов CiteSeer, CiteULike, Google Scholar, arXiv.org, PubMed и многих других.

1.4 Структура метаданных научных публикаций и их категоризация

В большинстве случаев, наукометрические базы данных не содержат в себе полного содержимого научной публикации, а только информацию о ней и ссылку на исходный документ. Эту информацию называют метаданными публикации. Метаданные – это данные о данных или структурированные данные, представляющие собой характеристики описываемых сущностей для целей их идентификации, поиска, оценки, управления ими, набор допустимых структурированных описаний, которые доступны в явном виде и предназначение которых может помочь найти объект. Термин используется в контексте поиска объектов, сущностей, ресурсов. На рис. 1.3 показан пример метаданных в наукометрической базе.

Извлечение информации из слабоструктурированных веб-страниц ; Вилучення інформації із слабоструктурованих веб-сторінок ; Extracting information from the semistructured web pages

	Open Access
Назва:	Извлечение информации из слабоструктурированных веб-страниц ; Вилучення інформації із слабоструктурованих веб-сторінок ; Extracting information from the semistructured web pages
Автор:	Коляда, Андрей Сергеевич ; Гогунский, Виктор Дмитриевич
Опис:	Рассматривается проблема обработки данных, представленных в публичном доступе глобальной сети интернет. Определены задачи, решение которых позволяет разрешить проблему. Представлен способ извлечения информации из слабоструктурированных веб-страниц на примере наукометрических баз данных. Разработано программное обеспечение для автоматизации процесса извлечения информации из наукометрических баз данных и сохранения их с возможностью дальнейшей обработки. ; Розглядається проблема обробки даних, представлених у публічному доступі глобальної мережі інтернет. Визначено завдання, рішення яких дозволяє вирішити проблему. Представлений спосіб вилучення інформації із слабоструктурованих веб-сторінок на прикладі наукометричних баз даних. Розроблено програмне забезпечення для автоматизації процесу вилучення інформації із наукометричних баз даних і збереження їх з можливістю подальшої обробки. ; In most cases, information that is given on the Internet open access has no strictly-defined structure. Web pages, i.e. information sources on the Internet, can have a non-uniform layout within a resource. When it concerns the processing from such sources, there appears a problem of extracting useful information from semistructured data. The method for solving the problem involving the so-called "web scraping" approach is proposed in the paper. The essence lies in simulating a human operation on a web resource using a lowlevel Hypertext Transfer Protocol (HTTP). This approach makes possible working with any data structures which become known after a preliminary data source analysis. The examples of extracting information from web pages of scientometric of a subsequent result has been developed. Further studies include the possibility of intelligent processing of extracted information to filter out irrelevant data
Видавництво:	PRIVAT COMPANY "TECHNOLOGY CENTER"
Рік видання:	2014-02-05
Ресурси:	Східно-Європейський журнал передових технологій; Том 1, № 9(67) (2014); Інформаційно-керуючі системи; 51-54 ; Восточно-Европейский журнал передовых технологий; Том 1, № 9(67) (2014); Информационно-керуючі системи; 51-54 ; Eastern-European Journal of Enterprise Technologies; Том 1, № 9(67) (2014); Інформаційно-керуючі системи; 51-54
Тип документа:	info:eu-repo/semantics/article ; info:eu-repo/semantics/publishedVersion
Мова:	rus
Предмет:	веб-страница; интернет; информация; слабоструктурированный; извлечение ; УДК 004.62 ; веб-сторінка; інтернет; інформація; слабо структуриований; вилучення ; УДК 004.62 ; web page; internet; information; semistructured; extraction ; УДК 004.62
Авторське право:	Author (authors) guarantee(s) thatis(s) exclusive copyrights on this page provided to the editors of "Eastern-European journal of enterprise technologies" (ISSN 1729-3774). Author(s)(-s) the right to publish editorial, printing and other circulation issues. In addition, each copy of the journal, which published the article must contain the name of the author. Author(s)(-s) the right to distribute publications editorial and article metadata and bring them to the public, processing and systematization, and inclusion in various databases and information systems, an area that is not limited. The authors give their consent to the collection and processing of personal data for inclusion in the database according to the Law of Ukraine No.2297-УІ "О защите персональных данных" from 01.06.2010. ; Наше видання використовує положення про авторські права CREATIVE COMMONS для журналів відкритого доступу. Автори, які публікуються у цьому журналі, погоджуються з наступними умовами: 1. Автори залишають за собою право на авторство своєї роботи та передають журналу право першої публікації цієї роботи на умовах ліцензії Creative Commons Attribution License, котра дозволяє іншим особам вільно розповсюджувати опубліковану роботу з обов'язковим посиланням на авторів оригінальної роботи та першу публікацію роботи у цьому журналі. 2. Автори мають право укладати самостійні додаткові угоди щодо неексклюзивного розповсюдження роботи у тому вигляді, в якому вона була опублікована цим журналом (наприклад, розміщувати роботу в електронному сховищі установи або публікувати у складі монографії), за умови збереження посилання на першу публікацію роботи у цьому журналі. ; Наше видання використовує положення об авторських правах CREATIVE COMMONS для журналів відкритого доступу. Автори, які публікуються в цьому журналі, погоджуються з наступними умовами: 1. Автори залишають за собою право на авторство своєї роботи та передають журналу право першої публікації цієї роботи на умовах ліцензії Creative Commons Attribution License, котра дозволяє другим особам вільно розповсюджувати опубліковану роботу з обов'язковою посиланням на авторів оригінальної роботи та першу публікацію роботи у цьому журналі. 2. Автори мають право укладати самостійні додаткові угоди, які стосуються неексклюзивного розповсюдження роботи у тому вигляді, в якому вона була опублікована цим журналом (наприклад, розміщувати роботу в електронному сховищі установи або публікувати в складі монографії), за умови збереження посилання на першу публікацію роботи у цьому журналі.
Зв'язки:	http://journals.urau.ua/eejet/article/view/19496/18777
URL:	http://journals.urau.ua/eejet/article/view/19496
Контент-провайдер:	Scientific Periodicals of Ukraine (Ukrainian Research and Academic Network) / Наукова періодика України

Рисунок 1.3 – Метаданные публикации в наукометрической базе Base-search

Метаданные научных публикаций предоставляют не только наукометрические базы данных, но и другие источники, например, издатель. На

рис. 1.4 показан пример метаданных публикации от издателя. Как видно из примеров, метаданные состоят из набора пар «ключ – значение». Поле «ключ» указывает на тип информации содержащейся в поле «значение». Набор метаданных может быть разным для каждого источника, так как в основном эта информация предоставляется для ознакомления пользователя и не ориентирована на обработку их автоматизированными средствами.

Розвиток системної концепції наукометричних баз даних

Заголовок (російською):
Развитие системной концепции наукометрических баз данных

Заголовок (англійською):
The development of the system concept of scientometric databases

Автор(и):
V.D. Gogunsky
V.O. Iakovenko
A.S. Kolyada

Ключові слова (укр):
науко-метричні бази даних; інформаційно-аналітична система; латентно-семантичний аналіз; аналізатор даних

Ключові слова (рус):
научно-метрические базы данных; информационно-аналитическая система; латентно-семантический анализ; анализатор данных

Ключові слова (англ):
scientometric databases; information-analytical system; latent semantic analysis; data analyzer

Анотація (укр):
Проаналізовано принципи роботи наявних науко-метричних баз даних. Запропоновано концепцію автоматизованої інформаційно-аналітичної системи для моніторингу інформації щодо публікацій науковців з України у міжнародних науко-метричних базах даних. Описано проблему розподілення отриманих статей між авторами та наведено її рішення.

Анотація (рус):
Проанализированы принципы работы существующих научно-метрических баз данных. Предложена концепция автоматизированной информационно-аналитической системы для мониторинга информации относительно публикаций ученых из Украины в международных научно-метрических базах данных. Описана проблема распределения полученных статей между авторами и приведено ее решение.

Анотація (англ):
The article shows an advantage of developing an automated system for monitoring scientific publications of Ukraine in the international scientific-metric databases. The main problems of this system are distribution of articles by author and authors with the same surname, name and patronymic (SNP). It is proposed to use data analyzer to solve these problems. Analyzer will get as an input a list of articles from international scientometric databases according to the specified SNP, and will return adjusted list of publications of authors without the same SNP. Latent semantic analysis is used in an automatic mode to improve system performance and to reduce user interaction with the analyzer. The whole information-analytical system's structure is tightly coupled, but at the same time each of its components executes functions specific only to it relying on the output from other components. The developed system will help to study the structure and evolution of the various branches of science in Ukraine.


Публікатор:
Київський національний університет будівництва і архітектури

Назва журналу, номер, рік випуску (рус):
Управление развитием сложных систем, номер 20, 2014

Назва журналу, номер, рік випуску (англ):
Management of Development of Complex Systems, Number 20, 2014

Мова статті:
English

Формат документа:
application/pdf

Документ:
 28.pdf

Дата публікації:
28 Декабрь 2014

Номер збірника:
20

Розділ:
ИНФОРМАТИЗАЦИЯ ВИШОЇ ОСВІТИ

Університет автора:
1 Odessa National Polytechnic University, Odessa; 2 Company "Tobii Technology" AB, Stockholm, Sweden

References:

1. Burkov, V. M., Beloschitsky, A. A., & Gogunsky, V. D. (2013). Options citation of scientific publications in scientometric databases. *Management of development of difficult systems. Kyiv, Ukraine: KNUCA*, 15, 134 - 139.

2. Gogunsky, V. D., Kolyada, A. S., & Iakovenko, V. O. (2014). *Scientometric data scientific publication "Management of development of difficult systems. Management of development of difficult systems. Kyiv, Ukraine: KNUCA*, 19, 6 – 11.

Рисунок 1.4 – Метаданні публікації журналу "Управление развитием сложных систем "

Но такие метаданные как название, автор(ы) и дата публикации являются основной информацией о публикации и содержатся практически во всех источниках предоставляемых их. Также нередко содержится информация об издательстве и ссылка (URL) на исходный документ публикации. Другие важные метаданные, такие как список использованных источников, тематика или направление исследования, аннотации, зачастую отсутствуют.

Однако на сегодняшний день существуют способы стандартизации метаданных публикаций для облегчения возможной обработки их автоматизированными средствами. Одним из них является использование специальных репозиториев, которые предназначены для документооборота определенного типа.

Среди таких репозиториев, которые связаны с публикационной активностью, можно выделить следующие:

- Eprints – пакет свободного программного обеспечения для построения архивов открытого доступа и в основном используется для создания коллективных архивов и научных журналов;

- Dspace – свободная платформа для институционального репозитория (для долгосрочного хранения цифровых материалов, используемых в академических исследованиях);

- Digital Commons – является открытым институциональным репозиторией и издательским решением, сочетающее традиционную функциональность с инструментами для рецензируемых публикации журнала, управление конференциями и мультимедиа;

- OJS (Open Journal Systems) – система предназначена для создания рецензируемых электронных журналов с открытым доступом и позволяет не только публиковать статьи в интернете, но и организовывать весь рабочий процесс издательского дела: приём, рецензирование и каталогизирование статей.

Такие репозитории, кроме своей основной функциональности, предоставляют возможность индексации метаданных содержащихся документов. Для веб приложений это реализуется с помощью добавления соответствующих мета тегов в html страницу. Существует несколько стандартов этих тегов, наиболее популярные и широко используемые среди них являются:

- Highwire Press теги (например, citation_title);
- Eprints теги (например, eprints.title);
- BE Press теги (например, bepress_citation_title);
- PRISM теги (например, prism.title);
- Dublin Core теги (например, DC.title) – устаревший формат, не рекомендованные к использованию.

Пример описания метаданных публикации с использованием Highwire Press тегов показан на рис. 1.5

```
<meta name="citation_title" content="The testis isoform of the phosphorylase kinase
  catalytic subunit (PhK-T) plays a critical role in regulation of glycogen
  mobilization in developing lung">
<meta name="citation_author" content="Liu, Li">
<meta name="citation_author" content="Rannels, Stephen R.">
<meta name="citation_author" content="Falconieri, Mary">
<meta name="citation_author" content="Phillips, Karen S.">
<meta name="citation_author" content="Wolpert, Ellen B.">
<meta name="citation_author" content="Weaver, Timothy E.">
<meta name="citation_publication_date" content="1996/05/17">
<meta name="citation_journal_title" content="Journal of Biological Chemistry">
<meta name="citation_volume" content="271">
<meta name="citation_issue" content="20">
<meta name="citation_firstpage" content="11761">
<meta name="citation_lastpage" content="11766">
<meta name="citation_pdf_url"
  content="http://www.example.com/content/271/20/11761.full.pdf">
```

Рисунок 1.5 – Метаданные научной публикации с использованием Highwire Press тегов

При использовании таких мета тегов можно автоматизировать процесс индексации и извлечения публикаций со страницы. А при использовании широкораспространенных стандартных тегов, можно предоставить эту возможность другим существующим системам. Примерами таких систем могут быть Академия Google и Mendeley. Обе системы поддерживают такие стандарты тегов как Highwire Press, Eprints, BE Press и другие.

Академия Google и Mendeley также предоставляют возможность связывания выбранных публикаций с Вашим аккаунтом (рис. 1.6). Таким образом, можно создавать репозитории с различным набором публикаций, например: Ваши публикации, интересные и/или полезные публикации, публикации по определенной тематике и другие.

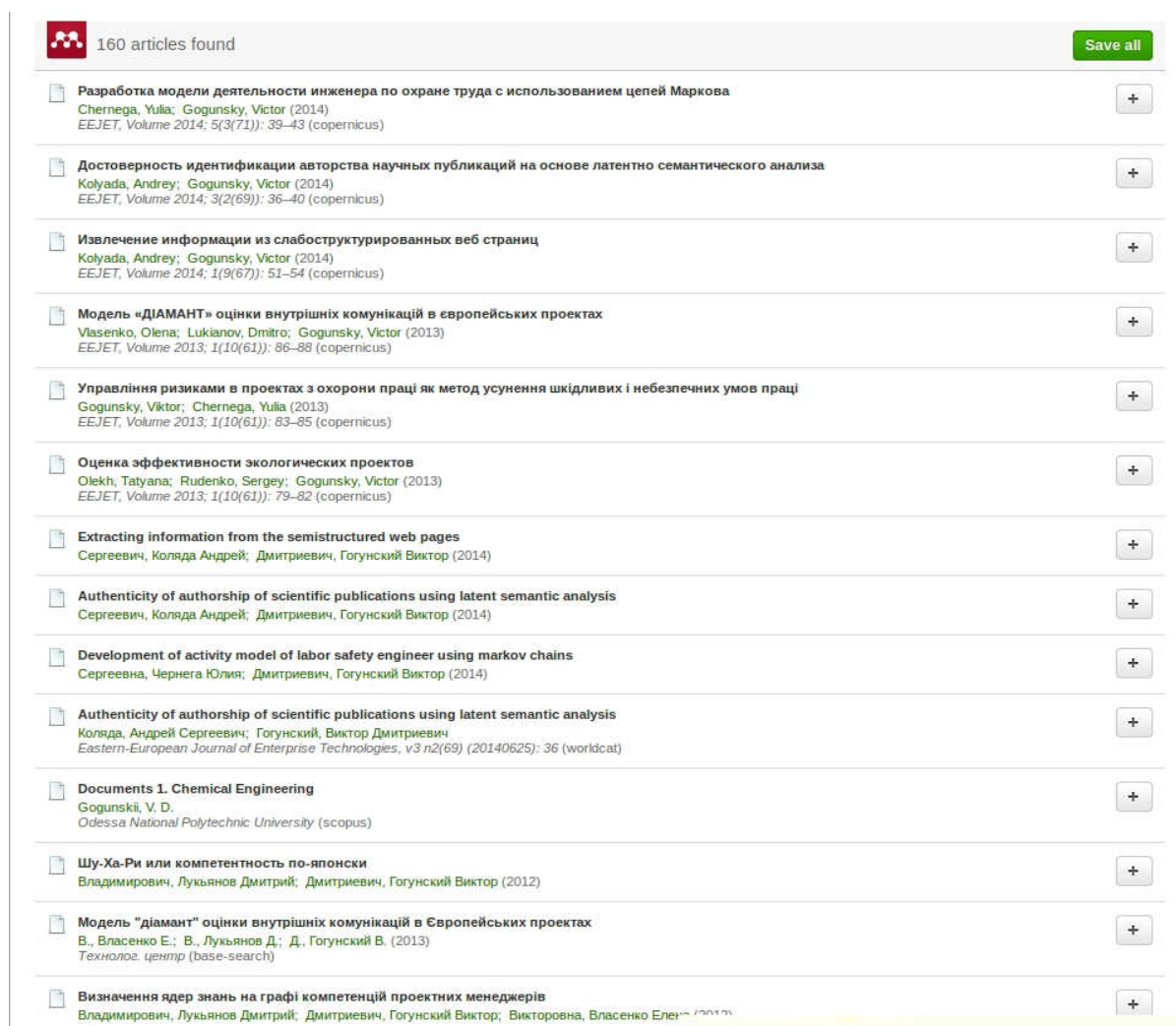


Рисунок 1.6 – Импорт публикаций в системе Mendeley

Для наполнения набора используется инструментарий выбранной системы. Например, Mendeley предоставляет пользователям букмарклет (небольшая JavaScript-программа, сохраняемая как браузерная закладка) для сохранения публикаций, находящихся на выбранной веб-странице.

Существует еще один способ связывания публикаций к конкретному человеку, точнее его уникальному идентификатору ORCID (англ. Open Researcher and Contributor ID) — единый международный реестр ученых (рис 1.7).

Рисунок 1.7 – Связывание публикаций в системе ORCID

В данном исследовании ведется схожая работа по связыванию научных публикаций, но автоматизируется процесс поиска и идентификации публикаций по различным источникам. Также поддерживается экспорт найденных публикаций в систему Mendeley, что позволяет создавать набор из своих публикаций одним щелчком.

1.5 Устройство поисковых систем

Поисковая система — это компьютерная система, предназначенная для поиска информации. Одно из наиболее известных применений поисковых систем — веб-сервисы для поиска текстовой или графической информации во Всемирной паутине. Основные составляющие поисковой системы: поисковый робот, индексатор, поисковик[75]. Как правило, системы работают поэтапно. Сначала поисковый робот получает контент, затем индексатор генерирует доступный для поиска индекс, и наконец, поисковик обеспечивает функциональность для поиска индексируемых данных. Чтобы обновить поисковую систему, этот цикл индексации выполняется повторно[75].

Поисковые системы работают, храня информацию о многих веб-страницах, которые они получают из HTML страниц. Поисковый робот или «краулер» (англ. Crawler) — программа, которая автоматически проходит по всем ссылкам, найденным на странице, и выделяет их. Краулер, основываясь на ссылках или исходя из заранее заданного списка адресов, осуществляет поиск новых документов, ещё не известных поисковой системе.

Поисковая система анализирует содержание каждой страницы для дальнейшего индексирования. Слова могут быть извлечены из заголовков, текста страницы или специальных полей — метатегов. Индексатор — это модуль, который анализирует страницу, предварительно разбив её на части, применяя собственные лексические и морфологические алгоритмы. Все элементы веб-страницы вычлняются и анализируются отдельно. Данные о веб-страницах хранятся в индексной базе данных для использования в последующих запросах. Индекс позволяет быстро находить информацию по запросу пользователя [76]. Поисковик работает с выходными файлами, полученными от индексатора. Поисковик принимает пользовательские запросы, обрабатывает их при помощи индекса и возвращает результаты поиска[75]. Когда пользователь вводит запрос в поисковую систему (обычно при помощи

ключевых слов), система проверяет свой индекс и выдаёт список наиболее подходящих веб-страниц (отсортированный по какому-либо критерию), обычно с краткой аннотацией, содержащей заголовок документа и иногда части текста[76].

Полезность поисковой системы зависит от релевантности найденных ею страниц. Хотя миллионы веб-страниц и могут включать некое слово или фразу, но одни из них могут быть более релевантны, популярны или авторитетны, чем другие. Большинство поисковых систем использует методы ранжирования, чтобы вывести в начало списка «лучшие» результаты. Поисковые системы решают, какие страницы более релевантны, и в каком порядке должны быть показаны результаты, по-разному[76].

Существует четыре типа поисковых систем: с поисковыми роботами, управляемые человеком, гибридные и мета-системы[77]:

– Системы, использующие поисковых роботов. Состоят из трёх частей: краулер («бот», «робот» или «паук»), индекс и программное обеспечение поисковой системы. Краулер используется для обхода сети и создания списков веб-страниц. Индекс – большой архив копий веб-страниц. Цель программного обеспечения – оценивать результаты поиска. Благодаря тому, что поисковый робот в этом механизме постоянно исследует сеть, информация в большей степени актуальна. Большинство современных поисковых систем являются системами данного типа.

– Системы, управляемые человеком (каталоги ресурсов). Эти поисковые системы получают списки веб-страниц. Каталог содержит адрес, заголовок и краткое описание сайта. Каталог ресурсов ищет результаты только из описаний страницы, представленных ему веб-мастерами. Достоинство каталогов в том, что все ресурсы проверяются вручную, следовательно, и качество контента будет лучше по сравнению с результатами, полученными системой первого типа автоматически. Но есть и недостаток – обновление данных каталогов выполняется вручную и может существенно отставать от

реального положения дел. Ранжирование страниц не может мгновенно меняться.

– Гибридные системы. Такие поисковые системы сочетают в себе функции систем, использующие поисковых роботов, и систем, управляемых человеком [81].

– Мета-системы. Метапоисковые системы объединяют и ранжируют результаты сразу нескольких поисковиков [82].

На рис. 1.8 показана структура стандартной поисковой системы с использованием роботов.

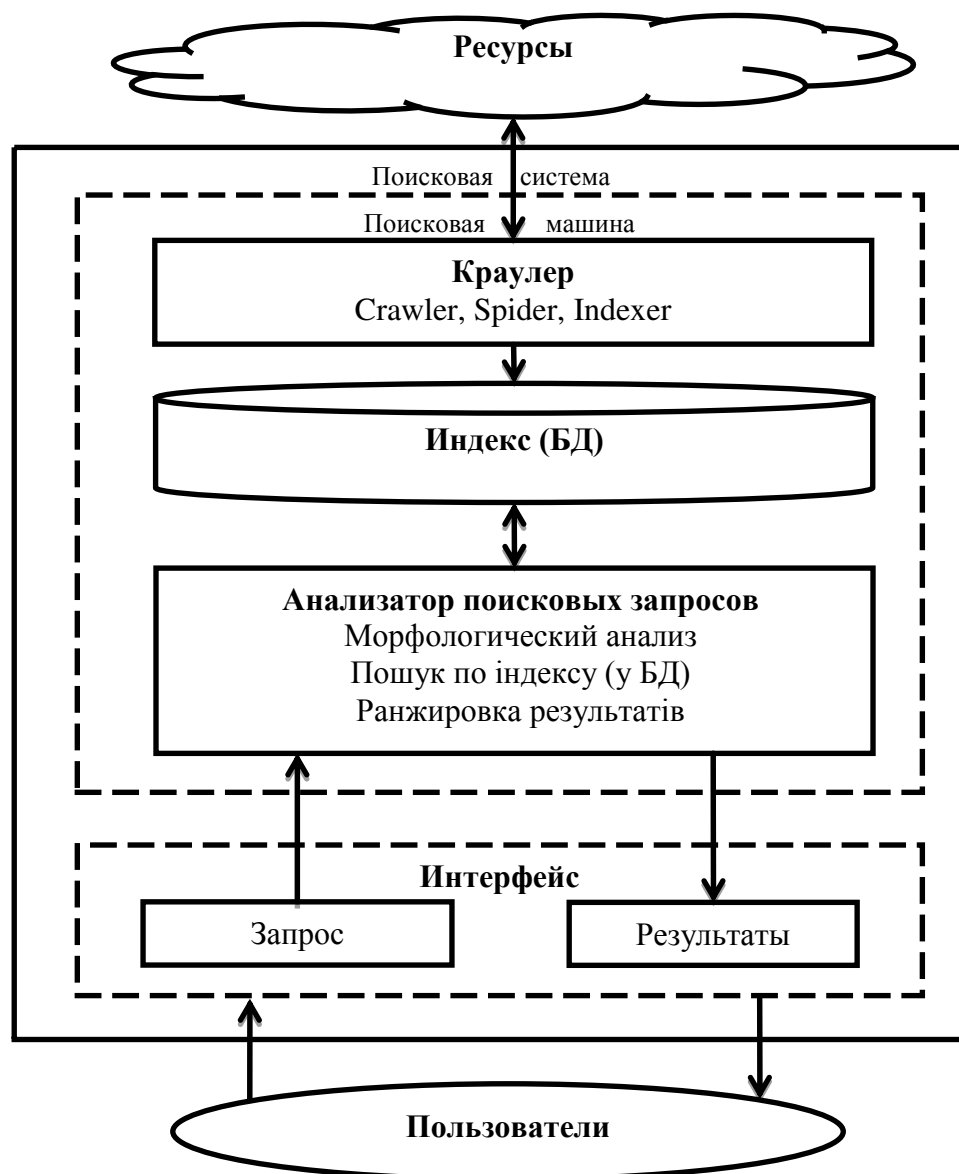


Рисунок 1.8 – Структура стандартной поисковой системы

1.6 Выводы

На основе анализа опубликованных работ и существующих программных продуктов и инструментов наукометрических баз установлено, что поиск публикаций в наукометрических базах данных, как правило, осуществляется только в пределах отдельных баз данных или репозиториев, что не позволяет определить интегральную оценку публикационной активности ученых; Наукометрические базы данных являются специализированными средствами автоматизации наукометрической деятельности. Появление их—стало следствием экспоненциального роста науки в середине XX века, когда научные сотрудники вынуждены были тратить почти половину своего рабочего времени на информационную деятельность. До недавнего времени о наукометрических базах знали и использовали их в основном библиотечные работники различных стран. Сегодня же, как минимум, каждый научный сотрудник Украины знает о них.

Наиболее известными и крупными наукометрическими базами данных являются Scopus и Web of Science. Среди некомерческих наукометрических баз данных, в которых индексируются научные публикации можно назвать следующие: Copernicus, BASE, DOAJ, Science Index, WorldCat, MLibrary.

Публикационная активность научных сотрудников является одним из основных факторов, который учитывается при определении мировых рейтингов высших учебных заведений (ВУЗ). Поэтому МОН Украины целенаправленно ориентирует публикационную деятельность научных сотрудников на вхождение в мировое научное сообщество. Несколько важных шагов было разработано в этом направлении: изменены требования к научным изданиям и к финансированию проектов научных исследований и разработок; ужесточены требования к планированию диссертационных исследований, формулировка их тематик; изменены критерии оценки ВУЗ Украины и другие. И заключающим этапом является создание инструментов «измерения» публикационной

активности ученых с последующим формированием информационно-аналитической системы мониторинга публикационной активности ученых ВУЗ Украины. С этим этапом и связана данная работа по извлечению метаданных публикаций из наукометрических баз данных.

Метаданные публикации это структурированные данные, представляющие характеристики публикации для целей их идентификации, поиска, оценки или управления. Для облегчения обработки их автоматизированными средствами, различные репозитории используют специальные мета теги на веб странице: Highwire Press теги, Eprints теги, BE Press теги и другие. Автоматизация извлечения метаданных публикаций из различных наукометрических баз данных позволит создавать средства мониторинга публикационной активности научного сотрудника.

Результаты раздела 1 опубликованы в работах автора [1, 2, 6-10].

РАЗДЕЛ 2

РАЗРАБОТКА МОДЕЛИ ИЗВЛЕЧЕНИЯ ИНФОРМАЦИИ ИЗ НАУКОМЕТРИЧЕСКИХ БАЗ ДАННЫХ

2.1 Веб интерфейс как основной доступ к информации из наукометрических баз данных.

Интерфейс – совокупность возможных способов и методов взаимодействия двух систем, устройств или программ для обмена информацией между ними с определёнными их характеристиками, а также характеристиками соединений, сигналов обмена и т. п.

В случае, если одна из взаимодействующих систем — человек, чаще говорят лишь о второй системе, то есть об интерфейсе той системы, с которой человек взаимодействует (пользовательский интерфейс). Одним из примеров пользовательского интерфейса является веб интерфейс программ во всемирной сети Интернет. Веб-интерфейс – это совокупность средств, при помощи которых пользователь взаимодействует с веб приложением.

На сегодня насчитывается значительное количество международных наукометрических баз данных, которые различаются структурой и способом хранения информации. Программный интерфейс для доступа к каждой базе, если и существует, то зачастую не афишируется. Не существует единого, универсального интерфейса, который подходил бы ко всем базам. Но есть один интерфейс, который имеют многие наукометрические базы данных и ориентирован больше на пользователя, нежели на программное обеспечение.

Доступ к содержимому (в ограниченном виде) предоставляет веб-интерфейс. Пользователь, с помощью веб-браузера, загружает веб-страницу определенной наукометрической базы данных и, используя поиск по заданным параметрам, получает необходимую информацию на странице.

В данном исследовании предлагается извлекать программным способом информацию, ориентированную на пользователя (человека). Таким образом, имитируется работа пользователя, который загрузил бы тысячи веб-страниц и собрал информацию определенной структуры в локальное место хранения.

Классическим и наиболее популярным методом создания веб-интерфейсов является использование HTML с применением CSS и JavaScript'a. Существует несколько техник для анализа и обработки содержимого веб-страницы:

- работа с веб-страницей как с обычным текстом – извлечения исходного кода страницы и применение утилит для работы с текстовой информацией (например, регулярные выражения);

- использование языков запросов для слабоструктурированных данных. Веб-страницы представлены языком разметки, который состоит из именованных тегов. Привязываясь к этим тегам, можно выполнять анализ и обработку данных (XPath – язык запросов к XML-подобным документам).

- построение и работа с объектной моделью документа (DOM) – не зависящий от платформы и языка программный интерфейс, позволяющий получить доступ к содержимому веб-документов, а также изменять содержимое, структуру и оформление их. Данный способ имитирует работу веб-браузера.

Развитие интернет-технологий в области организации хранилищ данных, репозиторий и электронных библиотек с предоставлением доступа к базам данных научных публикаций, создает условия для развития исследований в разных областях знаний, которые в определенной мере отображаются в научных публикациях. Именно множество публикаций составляет основу формирования новых знаний. Разработка интеллектуального интерфейса для взаимодействия с различными наукометрическими базами данных позволит существенно упростить поиск информации [1].

2.2 Модель извлечения информации из слабоструктурированных веб страниц.

Обработка содержимого веб страницы порождает задачу анализа и идентификации слабоструктурированных данных. Слабоструктурированные представления данных отличаются отсутствием строгих структур таблиц и отношений в моделях реляционных баз данных, тем не менее, эта форма данных содержит теги и другие маркеры для отделения семантических элементов, а также для обеспечения иерархической структуры записей и полей в наборах данных [35]. Проблема заключается в анализе информации содержащейся на веб-странице.

Глобальная сеть Интернет является самым большим источником данных, большая часть которых представляется в виде веб-страниц, не имеющих строго формализованной структуры. Извлечением информации из таких источников занимаются такие крупные корпорации как Google и Microsoft. Для качественного поиска используются сложные математические модели, семантический анализ и другие способы анализа информации. Поэтому данные, которые поступают на вход этих систем, должны быть структурированы определенным образом. Однако большинство наукометрических баз данных представлены в форме уникальных структур, что усложняет получение структурированных данных из подобных слабоструктурированных веб страниц.

Извлечение структурированных данных из веб страниц сводится к решению следующих задач [36 – 39]:

- поиску и получению целевых страниц для извлечения данных (проблема навигации);
- распознаванию участков, содержащих нужные данные (проблема распознавания данных);
- поиску структуры найденных данных (проблема поиска общей структуры данных);

- обеспечению однородности извлекаемых данных (проблема сопоставления атрибутов извлекаемых данных);
- объединению данных из разных источников (проблема объединения данных).

Для решения задачи извлечения данных на примере наукометрических баз данных [11], предлагается модель программного обеспечения (рис. 2.1), которая состоит из следующих компонентов:

- программа для извлечения данных из конкретных НМДБ;
- блок фильтров извлеченных результатов;
- база данных для конечных результатов.

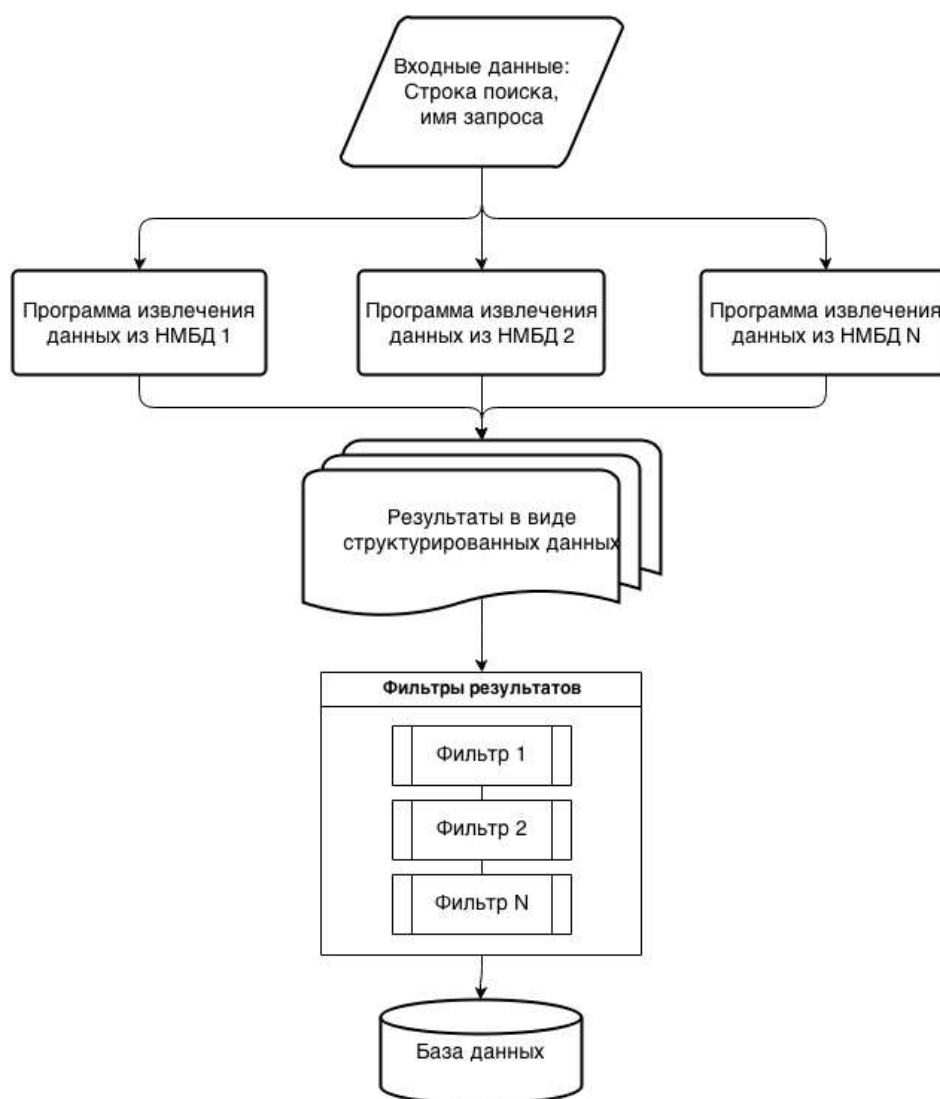


Рисунок 2.1 – Схема работы программного комплекса извлечения информации из наукометрических баз данных

Для каждой НБД создается отдельная программа извлечения данных, так как все базы имеют разный интерфейс и структуру. Эти программы содержат в себе логику работы с конкретной НБД, а также необходимые параметры, константные данные для выполнения этой работы.

После завершения работы программ извлечения данных выходные результаты каждой из них собираются в общий массив, который далее передается в блок фильтров. Блок состоит из одного или нескольких фильтров, которые отбрасывают нерелевантные результаты, согласно некоторым параметрам, специфичным для этого фильтра. Например, результаты программ извлечения данных могут содержать записи не соответствующие запросу поиска. Для этого можно использовать фильтр, который будет оставлять, только результаты соответствующие поисковой строке. Также можно использовать фильтр для отбрасывания результатов однофамильцев, удаления дубликатов и т.д.

После обработки блоком фильтров, оставшийся набор результатов записывается в базу данных для дальнейшего представления и анализа.

Рассмотрим подробнее работу программы извлечения данных. На рис. 2.2 показана общая схема ее работы. Каждая программа может иметь различия в деталях из-за слабоструктурированных данных.

На вход программы подается строка поиска. Далее выполняется запрос на загрузку начальной страницы конкретной НБД, где программным путем имитируется работа пользователя браузера - вводится строка поиска в поле поиска и выполняется запрос на выдачу результатов. Из веб-страницы результатов, с помощью подпрограммы, извлекаются все поля каждой записи и преобразовываются в структуру строго определенную программным комплексом. Отсутствующие поля остаются пустыми. Таким же образом извлекаются ссылки на другие страницы, так как большое количество результатов может быть разделено на страницы. Далее выполняются запросы на эти ссылки, и процесс повторяется сначала, пока не будут обработаны все

результаты либо сработает ограничение на количество, которое для каждой НБД устанавливается отдельно (один из параметров). В конце работы программы извлечения данных получаем набор структурированных данных, готовых для дальнейшей обработки.

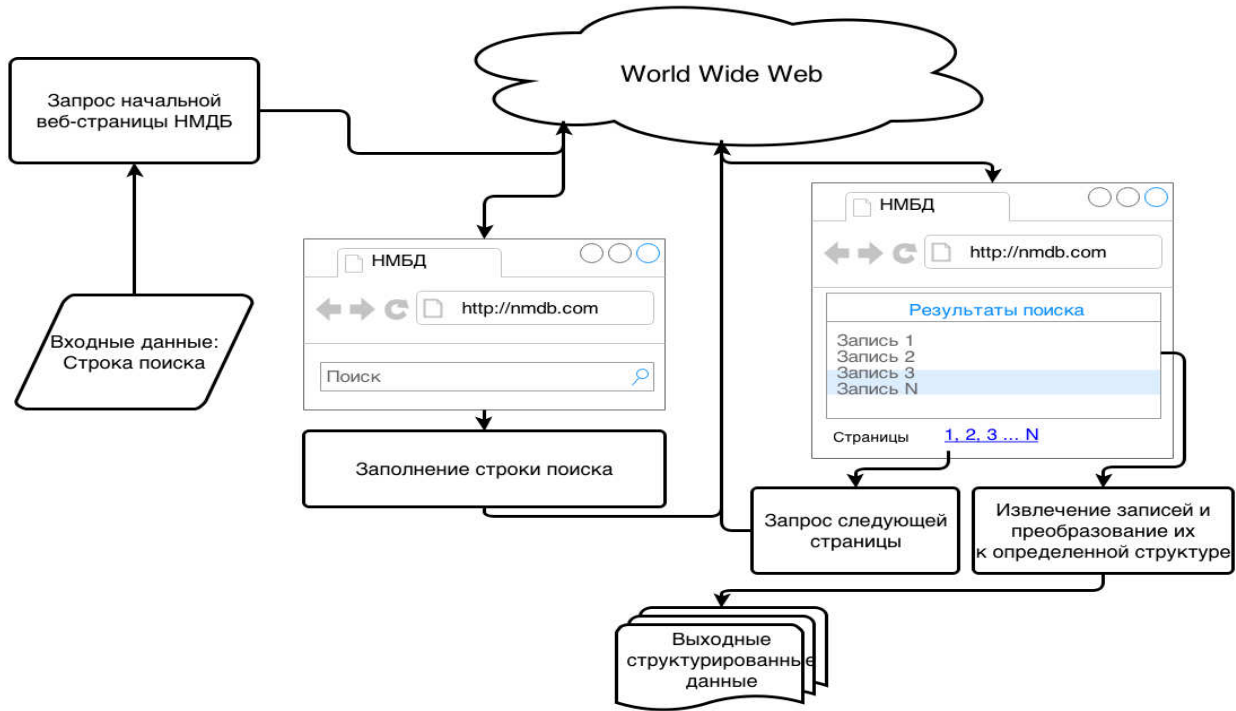


Рисунок 2.2 – Общая схема работы программы извлечения данных из НМДБ

Рассмотрим сам процесс извлечения информации со страницы НБД. Веб-страница, которую возвращает сервер, отформатирована с использованием языка разметки (в основном HTML), для дальнейшего отображения в том или ином виде с помощью специальной программы (веб-браузер). На рис. 2.3 показан пример визуализации веб-браузером некоторой области данных и исходный код этих данных. Здесь, например, название статьи “Features of Digital Devices Design of Modern PLD of the Xilinx Incorporation” заключено в следующие специальные последовательности символов, называемые тегами: `<h1 class="Title">Здесь название статьи</h1>`. Для извлечения этой информации, выполняется поиск этих тегов и извлекается их содержимое. Таким образом, заполняется одно из полей результатов. Для автоматизации

этого процесса, программы извлечения данных используют язык запросов к элементам языка разметки (Xpath).



Рисунок 2.3 – Данные с веб-страницы и их исходный код языка разметки HTML

Исследуем особенности извлечения данных из наукометрических баз данных, поддерживаемых разработанным программным обеспечением. На текущий момент определена следующая структура данных для каждой публикации (табл. 2.1).

Таблица 2.1 – Структура извлеченной информации

Поле	Описание
Наукометрическая база	Название базы источника публикации
Авторы	Авторы публикации
Название	Название публикации
Дата	Дата публикации
Источник	Источник публикации или издательство
Описание	Аннотация или краткое описание публикации
URL	Веб-ссылка на публикацию

BASE (base-search.net)

Наукометрическая база данных BASE позволяет выполнять поиск на разных языках и не задает строгих правил задания строки поиска (например, инициалы автора могут быть с точками или без них, а также слитно).

Результаты поиска подаются в визуальном структурированном виде, имеются следующие поля: название публикации, автор(ы), предмет,

издательство, год и URL. Но исходный код на языке разметки HTML имеет сложную структуру и вдобавок имена тегов зависят от языка интерфейса сайта. Поэтому перед началом работы с этой базой, мы устанавливаем язык интерфейса – английский. Результаты поиска находятся внутри тегов-контейнеров с именем класса “ResultsContent”. Для каждого результата мы анализируем его содержимое: теги с именем класса “ItemLeft_en” содержат имя поля, а теги с именем класса “ItemRight_en” – значение. Далее адаптируем эту информацию под структуру данных (табл. 1) и получаем извлечённую запись.

Scopus (scopus.com)

Поиск по наукометрической базе данных Scopus выполняется только на латинице. При этом для фамилии и инициалов имеются два разных поля ввода. Инициалы следует указывать с точкой. Работа с этой базой данной имеет особенности, в основном, из-за того, что результаты поиска – это информация об авторе. Поэтому для Scopus окончательная структура данных расширена на 2 поля: количество документов и h-индекс.

Результаты поиска представляются в виде таблицы со следующими полями: автор(ы), количество документов, предмет и т.д. Если автор имеет ссылку на расширенную информацию, мы переходим по этой ссылке, а также записываем ее с соответствующее поле (URL). На странице расширенной информации данные представлены в виде таблицы, состоящей из трех колонок: имя поля, разделитель, значение поля. Перебирая строки таблицы, мы заполняем выходную структуру данных.

Science Index (elibrary.ru)

Поиск поддерживается на многих языках. Для более эффективного поиска по автору, используется расширенный поиск, где указывается фамилия автора и инициалы, которые разделены пробелом.

Результаты поиска подаются в таблице, каждая строка которой содержит неструктурированную информацию: название статьи, автор(ы), источник, URL

и дата публикации. Для адаптации этой информации под общую структуру, используются следующие манипуляции со строкой результата (рис. 2.4):

- название и URL публикации извлекается и тега ``, который находится внутри тега `<a>`;
- авторы публикации извлекаются из тега `<i>`, который находится внутри первого тега ``;
- из второго тега `` извлекается дата и источник публикации.

№	Публикация
1	<p>ПРАКТИЧЕСКИЙ ОПЫТ ПО ПРИВЛЕЧЕНИЮ ИНВЕСТОРА В ВЕНЧУРНОМ БИЗНЕСЕ <i>Палагин А.В.</i> Интеграл. 2008. № 6. С. 52-53.</p>

```

<a href="/item.asp?id=11708906">
  <b>
    ПРАКТИЧЕСКИЙ ОПЫТ ПО ПРИВЛЕЧЕНИЮ ИНВЕСТОРА В ВЕНЧУРНОМ БИЗНЕСЕ</b>
</a>
<br>
<font color="#00008f">
  <i>
    Палагин А.В.</i>
</font>
<br>
<font color="#00008f">
  <a href="/contents.asp?issueid=531610">Интеграл</a>
  "
  2008.
  
```

Рисунок 2.4 – Пример строки результата поиска в наукометрической базе данных Science Index

Mlibrary(lib.umich.edu)

Наукометрическая база данных Мичиганского университета Mlibrary предоставляет поиск на латинице и имеет расширенный режим поиска для задания интересующих параметров. Мы используем параметр “Автор” для поиска. Замечено, что запись инициалов через пробел выдает больше результатов.

Результаты поиска выдаются в виде списка с названием публикации и ссылкой на полное описание. Мы переходим по этим ссылкам и извлекаем информацию из содержимого. Структура исходной информации представлена в следующем виде: теги с именем класса “article-field-label” содержат имя поля,

теги с именем класса “article-field-value” – значение. Проходом по всем полям мы извлекаем информацию только интересующих нас (табл. 2.1).

WorldCat(worldcat.org)

Поиск по базе WorldCat также выполняется на латинице с использованием расширенного режима, где указываем параметр “Автор” и “Формат публикации – статья”. Как и с базой Mlibrary, инициалы автора в строке поиска следует указывать разделенные пробелом.

Результаты поиска – список публикаций с кратким описанием и ссылкой на полное описание. Снова мы переходим по всем ссылкам и работаем с информацией на этих страницах. Содержимое страниц этой наукометрической базы данных имеет хорошо выраженную структуру, что очень редко для веб-страниц. Здесь каждое поле имеет свой идентификатор, по которому мы извлекаем значение. Например, идентификатор “bib-author-cell” содержит значение поля “Авторы”, а “bib-publisher-cell” – значение поля “Издательство”. Таким образом, мы легко заполняем свою структуру данных (табл. 2.1).

2.3 Модель веб скрапинга для автоматизации извлечения данных из веб страниц.

Множество данных в слабоструктурированной системе всемирной паутины образует сложную структуру организации информационных взаимодействий, изменяющихся во времени. При этом некоторые издания могут быть включены в одну и более наукометрических баз. Число публикаций постоянно увеличивается. Форматы представления библиографических данных и в публикациях, и в наукометрических базах существенно отличаются. Поиск публикаций в такой разнородной неформализованной среде часто удается осуществить только в ручном режиме.

Процесс поиска в данной среде является более искусством, нежели информационной технологией и зависит от умений и навыков пользователя.

Проблема состоит в том, чтобы максимально формализовать и автоматизировать этот процесс.

Для решения этой проблемы нужен способ извлечения данных из наукометрических баз в структурированном виде для возможной дальнейшей их обработки.

Исследованиями в направлении извлечения информации из глобальной сети Интернет занимаются крупные компании Google, Yandex, Microsoft. Они используют результаты исследований в реализации поисковых машин, которые являются главным компонентом поисковых систем. Поисковая машина представляет собой комплекс программ, предназначенный для поиска информации.

Одной из главных функций поисковых машин является извлечение информации из сети. Далее происходит обработка результатов, их индексация для ускорения выдачи результатов поиска и повышения его релевантности.

Основными компонентами подсистемы сбора и извлечения информации являются:

- «Паук» (Spider) – программа для загрузки веб-страниц;
- «Краулер» (Crawler) – программа для автоматического прохода по всем ссылкам, найденным на странице.

Паук скачивает веб-страницы тем же способом, что и веб-браузер, то есть имитируется действие пользователя. Но веб-браузер отображает эту информацию в графическом виде, а паук сохраняет ее для дальнейшей обработки. Краулер выделяет все ссылки, присутствующие на странице и переходит по всем или определенным ссылкам, исходя из заданных заранее условий. Следуя по найденным ссылкам, он перенаправляет страницы пауку для их загрузки.

Робот Googlebot – это разработанная Google программа сканирования Интернета («паук»). Сканирование является процессом, в ходе которого робот Googlebot обнаруживает новые и обновленные страницы для добавления в

индекс. Google использует огромную сеть компьютеров, чтобы извлечь содержание миллиардов веб-страниц. Робот Googlebot функционирует автономно и применяет алгоритмический процесс: компьютерные программы определяют сайты, которые нужно сканировать, а также частоту сканирования и количество извлекаемых страниц на каждом сайте.

Процедура сканирования начинается с получения списка URL веб-страниц, который создается на основе результатов предыдущих сеансов сканирования. Его дополняют данные из файлов Sitemap, предоставленных веб-мастером. Просматривая эти сайты, робот Googlebot находит на каждой странице ссылки и добавляет их в список страниц, подлежащих сканированию. Все новые и обновившиеся сайты, а также неработающие ссылки помечаются для обновления в индексе.

Основная цель – разработать способ извлечения данных о публикациях по параметру *Автор* из наиболее известных наукометрических баз данных с возможностью расширения поддерживаемых источников. Второстепенной задачей является знакомство с наиболее известными наукометрическими базами данных.

Для автоматического поиска и извлечения данных используется подход, основанный на применяемом в поисковых машинах – веб-скрапинг.

Веб-скрапинг – это процесс извлечения информации из веб-страниц, который фокусируется на преобразовании неструктурированных данных в сети (например, в формате HTML) в структурированный формат данных, который может быть проанализирован и сохранен. Веб-скрапинг также относится к автоматизации работы во всемирной паутине. На рис. 2.5 схематически изображен этот процесс.

Веб-скрапинг, также использует программы типа паук и краулер для обхода и загрузки веб-страниц. В отличие от поисковых машин, сканируется узкий круг веб-страниц, заданный начальными условиями и извлекается только полезная информация.

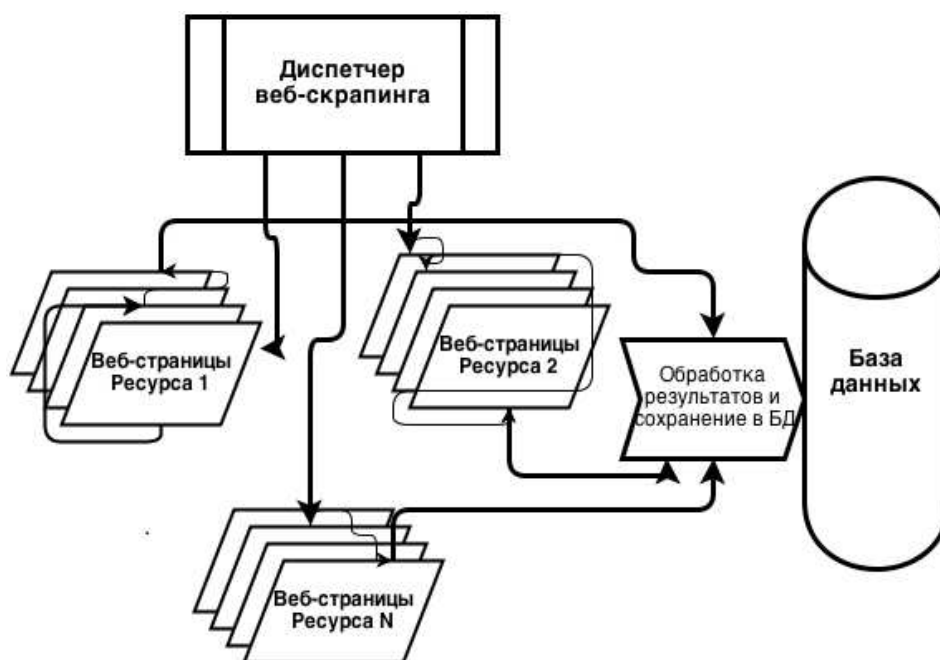


Рисунок 2.5 – Процесс веб-скрапинга

После извлечения информации в структурированном виде возможна дальнейшая ее обработка, которая может включать в себя фильтрацию результатов по некоторым критериям, подсчет различных коэффициентов и показателей, а также наиболее важная и сложная задача – определение однофамильцев и, соответственно, повышение точности результатов.

Исходя из содержимого извлеченной информации, можно решать проблему однофамильцев несколькими способами либо их комбинацией:

- семантический анализ темы или направления публикации;
- анализ ключевых слов публикации;
- анализ источника публикации.

В результате данного исследования спроектирована система извлечения информации о научных публикациях по параметру поиска *Автор*. Используя это свойство, программа выполняет поиск по известным ей наукометрическим базам данных и загружает результаты и извлекает информацию определенной структуры.

На данный момент поддерживаются следующие широко известные международные наукометрические базы данных:

- Scopus – библиографическая и реферативная база данных и инструмент

для отслеживания цитируемости статей, опубликованных в научных изданиях. Позиционируется издательской корпорацией Elsevier как крупнейшая в мире универсальная реферативная база данных с возможностями отслеживания научной цитируемости публикаций [56];

– Российский индекс научного цитирования (РИНЦ) – библиографическая база данных научных публикаций проекта eLIBRARY. РИНЦ выполняет функцию не только инструмента для оценки учёных или научных организаций на основе цитирования, но и авторитетного источника библиографической информации по научной периодике [58];

– BASE (Bielefeld Academic Search Engine) – многопрофильная поисковая система для научных интернет-ресурсов, созданная библиотекой университета Билефельд, Германия. Является одной из самых обширных поисковых систем публикаций в мире, особенно для открытого академического доступа к веб-ресурсам [59];

– Index Copernicus – интерактивная база данных из внесенной пользователем информации о ученом профиле, научных учреждений, публикаций и т.д. База данных имеет несколько инструментов оценки производительности, которые позволяют отслеживать влияние научных работ и публикаций, отдельных ученых или научно-исследовательских учреждений. Также Index Copernicus предлагает традиционное реферирование и индексирование научных публикаций. [60];

– Springer – международная издательская компания, специализирующаяся на издании академических журналов и книг по естественно-научным направлениям (теоретическая наука, медицина, экономика, инженерное дело, архитектура, строительство и транспорт). Является вторым по величине издательством в мире после Elsevier в области «STM» (science, technology, medicine — *англ.* наука, технологии, медицина) [61].

Ниже рассмотрены используемые технологии и средства, использованные при реализации системы извлечения неформализованной информации из веб-страниц.

Используемый формат извлеченных данных – текстовый формат обмена данными (JSON). Структура состоит из нескольких полей, таких как *Автор*, *Название* (публикации), *Источник*, *Дата*, *Наукометрическая база* и др. Структура не жесткая, может различаться набором полей для разных результатов, но, такие поля, как *Автор* и *Наукометрическая БД* являются обязательными.

Структура данных имеет следующий вид:

```
{ "title" : "Informational Model of Natural Language Processing",
  "url" : "http://hdl.handle.net/10525/263",
  "author" : [
    "Palagin, Aleksandr",
    "Gladun, Viktor",
    "Petrenko, Nikolay",
    "Velychko, Vitalii",
    "Sevruk, Aleksey",
    "Mikhailyuk, Andrey"
  ],
  "spider" : "base-search",
  "source" : "Institute of Information Theories and Applications FOI
ITHEA",
  "date" : "2008",
  "desc" : "The formal model of natural language processing in knowledge-
based information systems is considered. The components realizing
functions of offered formal model are described."}
```

Реализация загрузки веб-страниц, навигация по ссылкам и извлечение данных из веб ресурсов производится при помощи веб-скрапинг фреймворка Scrapy [62].

Извлеченные данные сохраняются в NoSQL базе данных MongoDB [63], потому что они не имеют жестких связей, как в реляционных базах данных.

Фреймворк Scrapy предоставляет удобный способ расширения числа поддерживаемых наукометрических баз данных путем добавления новой программы-паука ориентированного на работу с веб-ресурсом конкретной базы данных.

Используемый набор технологий и программного обеспечения позволяют создать программный продукт по извлечению информации из неоднородных и неформализованных источников (таких как наукометрические базы) и преобразованию ее в структурированный вид с возможной дальнейшей обработкой. Эти данные необходимы в первую очередь аспирантам и соискателям при подготовке к защите диссертаций [11, 22, 25]. Кроме того предлагаемая система может быть полезна при оценке деятельности ВУЗов [80].

Размещение публикаций в международных наукометрических базах может иметь позитивные последствия для науки Украины. На примере базы Scopus на сайте Национальной библиотеки Украины имени В. И. Вернадского показано, какую полезную информацию можно извлечь из нее: рейтинг ученых Украины, рейтинг организаций Национальной академии наук Украины, рейтинг высших учебных заведений Украины и т.д.

Представленный способ извлечения информации из международных наукометрических баз данных является своего рода универсальным интерфейсом для программного доступа к их содержимому (хоть и ограниченному). Используется процесс веб-скрапинга для обработки и извлечения неформализованных данных с дальнейшим приведением их к нормальному виду. Для поиска своих публикаций, автору требуется ввести свою фамилию и/или имя и запустить работу программы. Далее результаты в структурированном виде сохраняются в локальную (относительно наукометрических баз) базу данных и готовы к дальнейшей обработке или просмотру [40, 41].

2.4 Трудности извлечения данных из веб страниц и способы их решения

Описанная модель извлечения информации опускает сам процесс загрузки веб страниц из удаленного сервера. На практике этот процесс может быть нетривиальным, что требует дополнительной обработки. Вкратце, процесс получения веб страницы состоит из следующих шагов:

1. Клиент посылает запрос веб серверу.
2. Веб сервер, в качестве результата работы, генерирует веб страницу и отправляет ее клиенту.

В большинстве случаев, результирующая веб страница – это HTML страница, которую клиент обрабатывает и извлекает нужную информацию. Но также существуют веб страницы, которые кроме html разметки, содержат некоторый фрагмент программного кода, который клиент должен интерпретировать и выполнить, чтобы получить конечный результат. Такой программный код может содержать в себе обращение к серверу за дополнительной информацией или динамически создавать различные фрагменты HTML страницы. В связи с этим, обработка веб страницы в исходном виде от веб сервера усложняется. Информация может быть явно не представлена в результирующей веб странице, а генерироваться на этапе выполнения программного кода, который содержится на этой странице .

Еще одним препятствием для автоматического извлечения данных с использованием веб интерфейса является возможная блокировка доступа клиента к веб серверу. При этом веб браузер, который запущен с того же адреса клиента, может иметь доступ к запрашиваемой странице.

Таким образом, для достижения максимально возможных корректных результатов извлечения данных из веб страниц, нужно решить следующие задачи:

- обработка программного кода, который присутствует на веб странице;

– предоставить возможность идентификации клиента веб сервером, чтобы предотвратить блокировку доступа клиента к веб серверу.

Обе задачи связаны с тем, что информация из веб сервера запрашивается без помощи программы веб браузера, для которых, они в первую очередь предназначены. Но использование веб браузера усложняет автоматизацию процесса извлечения, требует зависимости от внешней программы, замедляет работу в целом, так как браузер может иметь лишний функционал, который не нужен для извлечения информации.

Максимально возможная имитация работы веб браузера для загрузки веб страницы и ее обработки (например, выполнение программного кода) поможет преодолеть указанные трудности. Проблема имитации работы веб браузера не нова, поэтому уже существуют ее решение – использование так называемых “безголовых” браузеров (англ. Headless browser).

“Безголовый” браузер – это браузер без графического интерфейса пользователя. Они обеспечивают автоматизированное управление веб-страницами в среде, аналогичной популярных веб-браузерах, но выполняются с помощью интерфейса командной строки или с помощью связи по сети. Они особенно полезны для тестирования веб-страниц, как они могут показывать и понимать HTML как и обычный браузер, в том числе расположение элементов, страницы, цвет, выбор шрифт, выполнение программного кода JavaScript и AJAX.

Наиболее распространенным “безголовым” браузером является PhantomJS – скриптовый браузер, который используется для автоматизации взаимодействия с веб-страницами. PhantomJS предоставляет программный интерфейс для использования его другими программами и основан на ядре Webkit, который используют такие браузеры как Safari и Google Chrome.

Используя такой браузер для загрузки веб страницы, дальнейший способ обработки информации, описанный в предыдущих пунктах этой главы, не нуждается в изменении. Таким образом, становится возможным извлечение

информации из различных веб страниц, в том числе динамически конструируемых на стороне клиента.

Таким образом, можно сделать вывод о том, что коллизии в основном могут возникать во время обработки динамических веб страниц из-за наличия исполняемого программного кода, который выполняется на стороне клиента. Для решения этой проблемы можно использовать так называемый «безголовый браузер», который загружает и формирует веб страницу как обычный браузер, но без графического интерфейса. При этом предоставляется программный доступ к текстовому содержанию страницы, что как раз и нужно для автоматизированной обработки.

2.5 Гибкое планирование при разработке программного обеспечения

На сегодняшний день, во всех сферах человеческой деятельности, происходит большое количество изменений за короткое время. Это конкурентная среда, где основным конкурентом являются мобильные, небольшие, динамично развивающиеся компании, которые открылась недавно, но уверенно завоевывает рынок благодаря реализации новых идей.

Согласно интервью с более чем 1100 руководителей из разных стран мира, предприятия будущего должны быть готовы к изменениям, разрушающими по своей природе старые догмы. Инновационные идеи должны опережать клиентские воображения и запросы. Развитие должно быть глобально комплексным и подлинным [68].

Все эти изменения требуют стать более адаптивными или даже повторного изобретения. Многие топ-менеджеры практически не имеют опыта работы в кризис. В исследовании, проведенном профессором Эдди Обенг (Henley Business School and Design Council) обнаружено, что более 70% руководителей высшего звена имеют очень мало опыта борьбы с такими условиями [68]. Обенг утверждал, что: «Мы развились в мире с того момента, когда мы могли учиться быстрее, чем среда окружения.» В результате этого

сдвига большинство понятий, передового опыта и предположений, обычно используемых для планирования, управления, и руководства являются устаревшими. Таким образом, одним из определяющих условий успешного развития является непрерывное совершенствование, поиск новых моделей и структур процессов, а не отдельных событий.

Поэтому исследования в области планирования проектов создания программных продуктов в условиях постоянных изменений являются актуальными. В качестве примера рассмотрим особенности реализации гибкой методологии разработки при создании программного обеспечения.

Анализ предыдущих исследований показывает, что индустрия программного обеспечения одна из первых ощутила на себе влияние изменений [69]. Один из фундаментальных уроков этой отрасли это гибкость – философия, которая воплощена в гибких методологиях разработки программного обеспечения.

Гибкие методологии ориентированы на частые поставки, итеративную разработку, и осуществляются при постоянном темпе, который измеряется и является прозрачным [70]. Люди и взаимодействие важнее процессов и инструментов, работающий продукт важнее исчерпывающей документации, сотрудничество с заказчиком важнее согласования условий контракта, готовность к изменениям важнее следования первоначальному плану [71].

Ниже приведены некоторые принципы, вокруг которых сконцентрировано планирование при использовании гибких методологий проектирования:

1. Идеи повсюду

Значимым двигателем инноваций является возрастание возможности взаимодействия между людьми - создателями бесконечного числа идей. Предприятия, работающие с применением гибких методологий, понимают это и создают среду (физическую и виртуальную) которая способствует идеям смешиваться, что дает следующие результаты:

- Связи между сотрудниками приносят новые мысли в организацию;

- Отношения рассматриваются как ценные активы;
- Компании меньше зависят от "запасов" знания, и больше связаны с "потокком" знаний [72].

2. Пользователи в центре

“Существует два пути роста компании – сконцентрироваться на том, что у вас хорошо получается или начать с клиента и работать исходя из его нужд, даже если это требует новых навыков” Джеф Безос, основатель Amazon.com [74].

3. Не единственное решение, а множество вариантов

Набор решений стоящих перед нами может быть ограничен решениями, которые кто-то сделал в прошлом, даже если обстоятельства поменялись. Вместо генерации одного решения одной проблемы, наша работа состояла должна состоять в создании такого количества опций для клиентов, насколько это возможно [69].

4. Проверь и учись

Провал бесполезен, пока не научишься у него. Продукт не должен быть безупречным. "Сделай что-то и изучи это, а не учи, потом делай". Гарет Кау [74].

5. Свобода планирования

Планы, которые установлены далеко наперед, являются жесткими и занимают огромное количество времени на планирование. Подробный план – враг приспособляемости [74].

Гибкие методологии подразумевают под собой не только итеративную разработку. Не менее важным считается гибкое планирование. Далее рассматриваются практические советы по планированию и управлению резервом работ, по организации итераций, а также измерениям и интерпретации производительности команды проекта.

Планирование итераций

Цель работы итерациями - налаживание регулярного взаимодействия со сторонами, заинтересованными в вашем продукте (клиентами, владельцем продукта и т.д.) Обычно длина итерации составляет две или четыре недели (иногда до шести недель). Короткие итерации проще планировать, так как быстрая оборачиваемость позволяет хорошо чувствовать сроки. Длинные итерации создают ложное ощущение, что для выполнения работы есть еще много времени, которое порой может быть губительным.

Резерв продукта

Владелец продукта отвечает за резерв продукта (product backlog, рис. 2.6) – набор пользовательских историй, которые должны быть реализованы в продукте. Пользовательские истории (user story) - это короткие (в одну строчку) предложения, описывающие требования клиентов. Пользовательские истории фокусируются на нуждах заказчика и не содержат в себе деталей реализации. Для каждой истории должен быть выставлен приоритет [70].

Резерв продукта может динамически изменяться. Можно добавлять новые истории, удалять или изменять приоритет существующих историй. С другой стороны, он должен быть не очень длинным. Нет никакой пользы в том, чтобы держать в очереди объем работы больший, чем вы можете выполнить за разумное количество времени. Если из резерва выпадает хорошая идея, то через некоторое время она туда вернется, если это действительно ценная идея, отвечающая нуждам заказчика.

Резерв релиза

Перед началом работы над очередным релизом продукта владелец продукта просматривает резерв продукта, выявляет истории, которые должны быть реализованы в этом релизе, и переносит их в резерв релиза (releasebacklog).

При создании резерва релиза все высокоуровневые пользовательские истории разбиваются на истории меньшего размера. Затем команда разработчиков оценивает в баллах каждую историю. Цель этого этапа – дать

команде разработчиков оценить новую функциональность, представленную небольшими порциями, и оценить сложность работы. Хотя резерв релиза – динамичный список, но после начала итерации владелец продукта уже не может изменять работу, запланированную на эту итерацию.

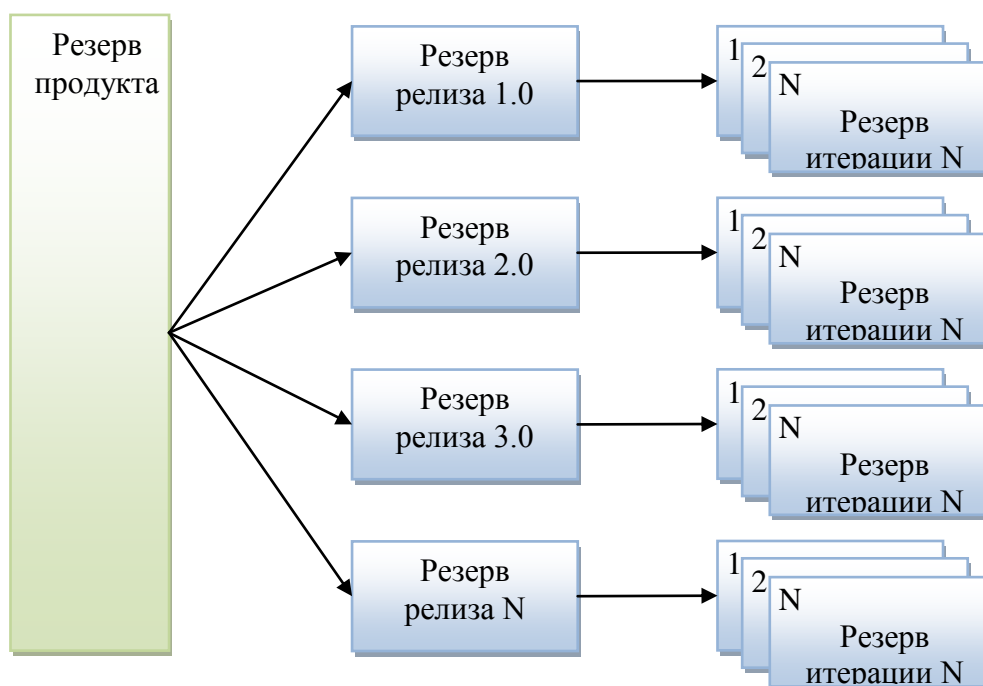


Рисунок 2.6 – Резерв продукта служит источником для заполнения остальных резервов

Резерв итерации

В начале итерации владелец продукта может скорректировать приоритеты историй в резерве релиза, чтобы команда разработчиков знала, чем ей заниматься. Половину первого дня итерации владелец продукта и команда разработчиков должны выделить для совещания по планированию итерации. На этом совещании команда разработчиков должна выбрать как можно больше высокоприоритетных историй и перенести их из резерва релиза в резерв итерации (iteration backlog). Для каждой пользовательской истории из резерва итерации команда выявляет задания, которые нужно выполнить для ее реализации, а также оценивает их. Необходимо выявить задания для всех, участвующих в работе над историей (разработчики, тестеры, технические писатели).

Оценка пользовательских историй

Команда разработчиков назначает баллы (story points) каждой пользовательской истории, а не единицы времени. Основной причиной является то, что на ранних этапах разработки вы не знаете, сколько времени вам фактически потребуется для реализации той или иной истории. Другой причиной использования баллов вместо единиц времени является то, что идеальное время для разных людей может быть разным. И, наконец, еще одна причина использования баллов: оценка идеального времени может быть ошибочно истолкована как оценка фактического времени. Фактическое время – это время, которое займет работа, с учетом всех прерываний рабочего процесса, происходящих в течение обычного рабочего дня. Например, на задание, оцененное в пять рабочих дней, фактически может уйти восемь-девять дней, если учесть ежедневные совещания, телефонные разговоры, ведение переписки по электронной почте и проверки.

Оценивая пользовательские истории в баллах, фактически сравнивается относительная сложность их реализации. Поначалу команде сложно оценивать истории в баллах ввиду отсутствия эталона для сравнения. Но со временем она будет делать все больше и больше оценок и постепенно их точность увеличится. Для оценки историй в баллах можно использовать способ Покера Планирования. Покер Планирования – это способ командного согласования оценки сложности или объема работы. Для каждого оцениваемого задания каждый член команды выбирает свою карту оценки (карты имеют значения 0, 0.5, 1, 2, 3, 5, 8, 13, 20, 40, 100), таким образом, что его карту не видит никто из других игроков. После того как все игроки сделали свой выбор, карты раскрываются (и происходит обсуждение).

Измерение производительности

Здесь производительность – это отслеживание в долгосрочной перспективе того, как много команда успевает сделать за итерацию. Точно так же, как и пользовательские истории, она измеряется в баллах. Чтобы заработать

баллы, назначенные какой-либо истории, команда должна завершить все задания этой истории в этой итерации.

Через несколько итераций работы и такого подсчета баллов, можно строить графики производительности, подобные тому, который показан на рис. 2.7. Польза подсчета производительности в том, что спустя некоторое время вы сможете использовать эту статистику для экстраполяции будущего графика работ.

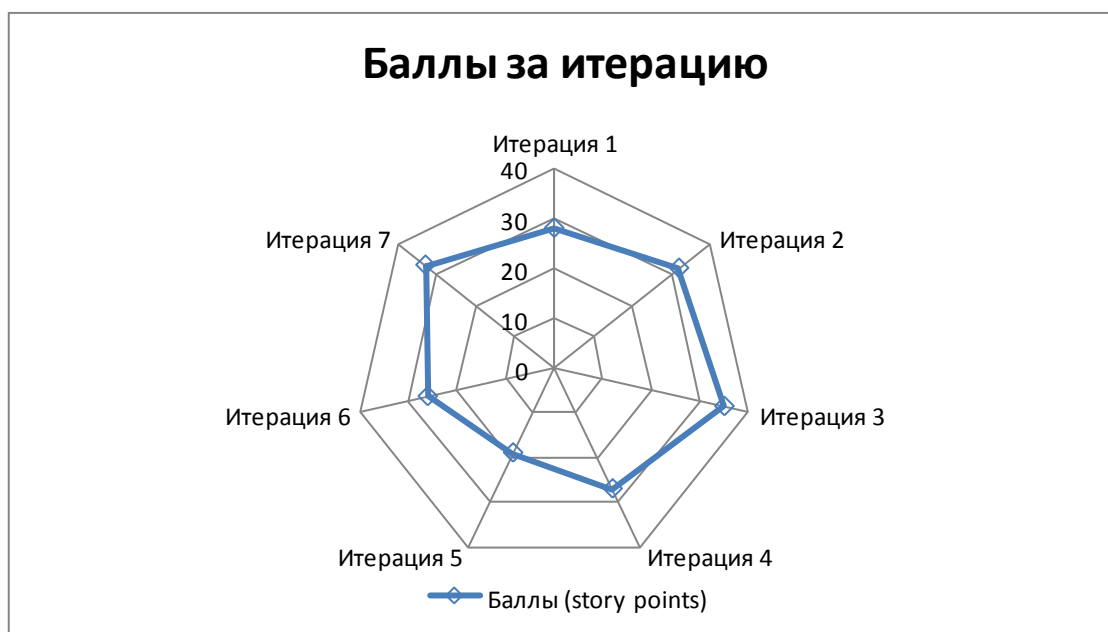


Рисунок 2.7 – График производительности на протяжении нескольких итераций.

На основе этих данных можно вычислить следующие показатели:

- Среднее количество баллов, зарабатываемое за итерацию – 28;
- Текущая производительность команды – 33;
- Средняя производительность за 3 самые медленные итерации – 24 баллов.
- Например, если до окончания работы над релизом остается 6 итераций, можно сделать следующие предположения:
 - При работе со средней производительностью команда сможет завершить историй еще на 168 баллов ($6 * 28$);

- При работе с текущей производительностью команда сможет завершить историй еще на 198 баллов ($6 * 33$);
- При работе с низкой производительностью команда сможет завершить историй еще на 144 балла ($6 * 24$).

Основная цель гибкого планирования – сделать как можно больший объем «известной» работы видимым всем участникам процесса. «Известной», потому что по мере получения новых знаний о выполняемой работе возможно добавление в резерв новых историй. Это позволяет владельцу продукта поддерживать приоритеты историй в актуальном состоянии и гарантировать, что работа постоянно будет идти над наиболее важными составляющими проекта вещами.

Ожидания конечного пользователя сдвигаются в режим реального времени. Компании разрабатывают совершенно новые отношения, как с работниками, так и с их клиентами. Жесткие, длительные циклы планирования могут удовлетворить компании, но они не годятся для клиентов. Маркетинг является адаптивным, актуальность определяется в данный момент. Гибкие методологии наряду с гибким планированием позволяют держать соответствующий ритм и быть конкурентоспособным.

2.6 Выводы

Почти каждая наукометрическая база данных имеет различную структуру и различный способ хранения информации. Единого программного интерфейса для автоматизированного доступа к ним не существует. Но многие предоставляют веб интерфейс в виде веб страниц для просмотра содержимого по заданному критерию (зачастую ФИО автора). Так как информация на веб странице в основном определена языком разметки HTML, существуют несколько способов обработки ее:

- Работа с веб страницей как набором символов (например, применение

регулярных выражений);

– Использование языков запросов для слабоструктурированных данных (XPath – язык запросов к xml подобным документам);

– Построение и работа с объектной моделью документа (DOM) – имитация работы веб браузера.

Использование перечисленных техник для извлечения данных из веб страниц называется веб скрапингом. Комбинация различных способов позволяет извлекать информацию практически из любой веб страницы, независимо от ее структуры и содержимого.

Для решения задачи извлечения данных из наукометрических баз данных предлагается модель программного обеспечения, которая состоит из так называемых «программ-пауков», которые обходят веб страницы определенной базы данных и извлекают метаданных публикаций по заданному критерию. Такой подход используют также поисковые машины (Google, Yandex), только в крупном масштабе – для сканирования всех страниц доступных в сети интернет. Трудности в основном возникают во время обработки динамических веб страниц, которые содержат программный код, который выполняется на стороне клиента. Для решения этой проблемы используется так называемый «безголовый браузер», который загружает и формирует веб страницу как обычный браузер, но без графического интерфейса. При этом предоставляется программный доступ к содержимому страницы, что как раз и нужно для автоматизированной обработки.

Исследования раздела 2 опубликованы в работах автора [1, 2, 18].

РАЗДЕЛ 3

ОБРАБОТКА И АНАЛИЗ ИЗВЛЕЧЕННЫХ ПУБЛИКАЦИЙ

3.1 Вероятностное тематическое моделирование

В настоящее время графические модели являются основным инструментом для построения вероятностных тематических моделей (probabilistic topic model). Модели со скрытыми переменными оказались особенно эффективными для выявления скрытых структур в текстовых коллекциях. Важным является подкласс ориентированных вероятностных тематических моделей (directed probabilistic topic models, DPTM), которые осуществляют мягкую кластеризацию и применяются для выявления тематики текстов в больших коллекциях документов. В терминах кластерного анализа тема (topic) — это результат би-кластеризации, то есть одновременной кластеризации и слов, и документов по их семантической близости. При мягкой кластеризации (soft clustering) каждое слово и каждый документ относится к нескольким темам одновременно с определёнными вероятностями. Таким образом, сжатое семантическое описание слова или документа представляет собой вероятностное распределение на множестве тем. Процесс нахождения этих распределений и называется тематическим моделированием [73].

Тематические модели активно развивались последнее десять лет. Предложено много моделей для решения задач моделирования текстовых коллекций в таких приложениях, как классификация документов, поиск похожих документов, поиск экспертов, выявление сообществ и анализ временных трендов.

Графические модели могут быть разделены на две основные категории: ориентированные и неориентированные графические модели. Эти типы можно далее разбить на параметрические и непараметрические.

Графические модели со скрытым тематическим слоем в последнее время успешно применялись в задачах поиска скрытых закономерностей в данных. Автоматическое выделение тематики текстов применялось для разбиения текстов по группам на основе семантической близости содержания. Эти модели позволяют классифицировать документы, но они ограничены предположением, что каждый документ относится только к одному кластеру. Модели мягкой кластеризации позволяют относить документ одновременно к нескольким кластерам, при этом каждый кластер ассоциируется с определённой темой, и каждый документ характеризуется оценками близости к каждой из тем. В ориентированных вероятностных тематических моделях (directed probabilistic topic model, DPTM) оценка близости документа к теме имеет вероятностный смысл и может интерпретироваться как доля содержимого документа, относящаяся к данной теме. DPTM — это относительно молодая область исследований в теории самообучения (обучения без учителя, *unsupervised learning*), представляющая в настоящее время значительный как теоретический, так и практический интерес. Одним из первых был предложен вероятностный латентный семантический анализ (*probabilistic latent semantic analysis, PLSA*, рис. 3.1), основанный на принципе максимума правдоподобия, как альтернатива классическим методам кластеризации, основанным на вычислении функций расстояния.

Вероятностный латентный семантический анализ (*probabilistic latent semantic analysis, PLSA*) — основан на введении слоя скрытых переменных для описания тематик в коллекции текстовых документов. Модель PLSA является важной вехой в развитии вероятностного моделирования текстов, и она, несомненно, полезна для задач информационного поиска. Однако она имеет довольно существенные ограничения. В PLSA каждый документ представляется числовым вектором, каждая компонента которого равна доле соответствующей темы в документе. Однако вероятностная модель не описывает ни закон распределения этих долей, ни вероятности самих

документов. В результате число параметров модели линейно растёт с ростом размера текстовой коллекции, что может приводить к переобучению. Кроме того, не понятно, как оценивать вероятности новых документов, не входивших в состав обучающей выборки.

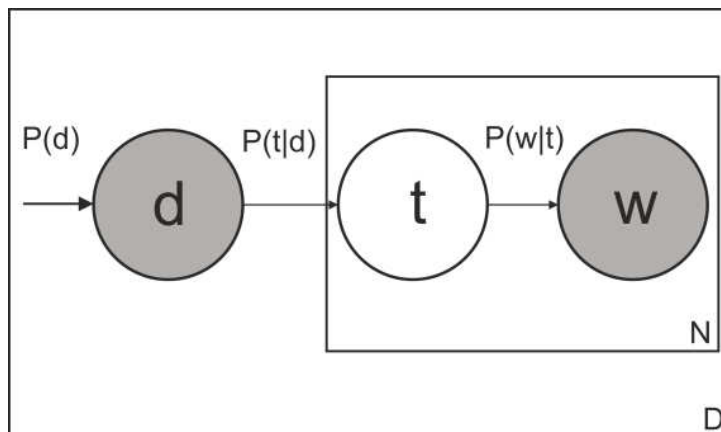


Рисунок 3.1 – Графическая вероятностная модель PLSA, где d – документ; w – слово; d , w – наблюдаемые переменные; t – тема (скрытая переменная); $p(d)$ – априорное распределение на множестве документов; $p(w|t)$, $p(t|d)$ – искомые условные распределения; D – коллекция документов; N – длина документа в словах.

Вслед за PLSA был предложен метод латентного размещения Дирихле (latent Dirichlet allocation, LDA) и его многочисленные обобщения. Применение тематических моделей позволяет получить ответ на целый ряд нетривиальных вопросов. Как выявлять смысл или тематику документов по их содержанию? Как осуществлять классификацию документов на основе этих скрытых тематических закономерностей? Как выявлять научные интересы авторов и находить экспертов в специальных областях знания? Как выявлять скрытые ассоциативные связи между отдельными исследователями или группами людей? Как подбирать коллаборации под проекты? Как выявлять тенденции в развитии научных направлений? Как выявлять роли людей в социальных сетях? Как осуществлять индексацию и автоматическое аннотирование документов?

Под параметрическими моделями будем понимать тематические модели, в которых темы изначально фиксированы и не меняются в процессе построения

модели. В непараметрических моделях число тем изначально не фиксировано, а сами темы настраиваются в процессе поиска наилучшего возможного модельного описания данных.

Ниже описываются концепции и терминология, лежащие в основе теории тематического моделирования.

Документ обычно состоит из множества слов, терминов (словосочетаний), специальных символов, таблиц, иллюстраций, и т.д. В исследованиях по тематическому моделированию типичными видами документов являются научные статьи и новостные сообщения.

Большую коллекцию текстовых документов принято называть *корпусом текстов* (text corpora). В исследованиях по тематическому моделированию часто используют общедоступные корпуса «NIPS proceedings» и «Cite seer». Оба содержат большое число научных статей и используются для тестирования различных методов обнаружения знаний. Известные корпуса «TREC AP» и «Reuter's» используются для тестирования методов анализа новостей.

В литературе по тематическим моделям понятие *темы* (topic) определяется по-разному, в зависимости от научной школы: «скрытые паттерны», «компактные описания смысла документов», «вероятностные (нечёткие) кластеры семантически связанных терминов», «связующее звено между терминами и другими объектами (документами, авторами, организациями, конференциями, и т.д.), которое позволяет находить скрытые ассоциативные связи между ними».

Формально тема определяется как дискретное (мультиномиальное) вероятностное распределение в пространстве слов заданного словаря. Документ может состоять из огромного числа слов, однако эти слова могут порождаться небольшим числом тем, как смесью мультиномиальных распределений.

Предположение о том, что для целей анализа текстов (в нашем случае — для выявления тематики) важна только частота слов, но не их порядок, называется *моделью мешка слов* (bag of words). Когда не важен порядок предложений в

документе или порядок документов в корпусе, соответствующим образом вводятся модели мешка предложений и мешка документов.

Основная идея тематического моделирования заключается в описании документа смесью тем, т.е. в определении документа как выборки слов, порождаемой смесью вероятностных распределений.

Тематическим моделированием называется решение обратной задачи. Каждый документ в корпусе текстов рассматривается как наблюдаемая случайная независимая выборка слов (мешок слов), порождённая некоторым, как правило небольшим, латентным подмножеством тем. По этим данным требуется восстановить вероятностные распределения всех тем в корпусе и определить, каким именно подмножеством тем порождён каждый документ.

3.2 Латентно-семантический анализ извлеченной информации с целью идентификации схожих публикаций

Проблема идентификации текстов естественного языка вычислительными машинами давно представляет научный интерес [40]. В настоящее время широко используются различные подходы распознавания речи, классификации текстовой информации, определение идентичности текстов [42]. Автоматизация извлечения информации из наукометрических баз данных сопряжена с необходимостью уточнения результатов запросов к базам данных в части исключения «двойников» авторов, у которых совпадают фамилии и инициалы [1, 25, 27]. Подобные проблемы возникают и при разработке систем автоматизированного обучения с открытыми тестами [43–46].

Одним из примеров извлечения смысла из текстов является поиск схожих документов или документов на определенную тематику. Стандартный поиск использует сравнение документов на наличие искомой строки или слов. Однако не всегда можно сформулировать точный запрос. Зачастую требуется поиск, который основан на анализе смысловой нагрузки документов. Одним из

подходов, который уже активно используют поисковые гиганты, является латентно-семантический анализ. Этот подход позволяет выявить закономерности в отношениях между понятиями и терминами в неструктурированной коллекции текстов.

Существует несколько способов смыслового анализа текстов, которые можно разделить на следующие группы [40]: лингвистический анализ; статистический анализ.

Первая группа ориентирована на определении смысла по семантической структуре текста и включает лексический, морфологический и синтаксический анализ. В настоящее время отсутствуют сложившиеся подходы к реализации задачи семантического анализа текстовой информации, что во многом обусловлено исключительной сложностью проблемы и недостаточно полной проработкой научного направления создания систем искусственного интеллекта.

Вторая группа – это, как правило, частотный анализ в тех или иных его вариациях. Суть анализа заключается в подсчете количества повторений слов в тексте и использовании результатов подсчета для конкретных целей. Всевозможные варианты различных реализаций подсчета слов и последующая обработка результатов подсчета образуют широкий спектр предлагаемых в данном классе методов и алгоритмов.

Одним из наиболее эффективных статистических подходов является латентно-семантический анализ (или латентно-семантическое индексирование) [48]. Авторы представили модель двухрежимного факторного анализа, которая основана на сингулярном разложении (SVD). Сингулярное разложение представляет термины и документы в виде векторов в пространстве выбираемой размерности, а скалярное произведение между точками пространства – их схожесть.

Латентно-семантический анализ начинается с построения матрицы документов и терминов – индексируемых слов [49]. Индексируемые слова это

слова, которые включаются в двух или более документах и имеют смысловую нагрузку (не являются предлогами, союзами и т.д.). Далее применяется сингулярное разложение этой матрицы на произведение трех матриц:

$$A = U \cdot S \cdot V^t, \quad (3.1)$$

где матрицы U и V – ортогональные, а S – диагональная матрица, на диагонали которой значения называются сингулярными значениями матрицы A .

Такое разложение обладает замечательной особенностью: если в матрице S оставить только k наибольших сингулярных значений, а в матрицах U и V – только соответствующие этим значениям столбцы, то произведение получившихся матриц S , U и V будет наилучшим приближением исходной матрицы A к матрице \hat{A} ранга k :

$$\hat{A} \approx A = U \cdot S \cdot V^t, \quad (3.2)$$

Основная идея латентно-семантического анализа состоит в том, что если в качестве матрицы A использовалась матрица индексируемые слова на документы, то матрица \hat{A} , содержащая только k первых линейно независимых компонент A , отражает основную структуру различных зависимостей, присутствующих в исходной матрице. Структура зависимостей определяется весовыми функциями индексируемых слов. Таким образом, каждое индексируемое слово терм и документ представляются при помощи векторов в общем пространстве размерности k . Близость между любой комбинацией индексируемых слов и/или документов легко вычисляется при помощи скалярного произведения векторов [50]. Как правило, выбор k зависит от поставленной задачи и подбирается эмпирически. Если выбранное значение слишком велико, то метод теряет свою мощьность и приближается по характеристикам к стандартным векторным методам. Слишком маленькое значение k не позволяет улавливать различия между похожими термами или документами.

Латентно-семантический анализ хорошо справляется с проблемой синонимии, но частично с проблемой полисемии, потому, что каждое слово определяется одной точкой в пространстве. Также этот анализ позволяет выполнять автоматическую категоризацию документов, основанную на их сходстве концептуального содержания. Также преимуществом латентно-семантического анализа является независимость от языка, так как это математический подход. Недостатком метода является снижение скорости вычисления при увеличении объема входных данных (например, при SVD-преобразовании).

Извлеченная информация из наукометрических баз данных нуждается в постобработке с целью определения схожих по смыслу публикаций, а также определения дубликатов. Применение латентно-семантического анализа для проекта по извлечению информации из наукометрических баз данных позволит разделить полученные публикации на категории с целью определения однофамильцев. Например, автор Иванов И. И. занимается исследованиями в области компьютерных наук, но результаты поиска его публикаций в наукометрических базах содержат много несоответствующих записей, так как есть еще один автор Иванов И. И., который опубликовал статьи по медицинской тематике. Латентно-семантический анализ позволит разделить публикации, которые относятся к разным концептам. Различные наукометрические базы могут содержать дубликаты публикаций в несколько измененной форме. Определение этих дубликатов также возможно с помощью латентно-семантического анализа.

Рассмотрим последовательность действий латентно-семантического анализа, изображенную на рис. 3.1, к некоторому набору документов. В качестве примера будет использоваться небольшой список названий публикаций, извлеченных из наукометрических баз данных для автора Колесниковой Е. В. (табл. 3.1). Выбор автора хорошо подходит для примера, так как существует несколько авторов с одинаковой фамилией и инициалами.

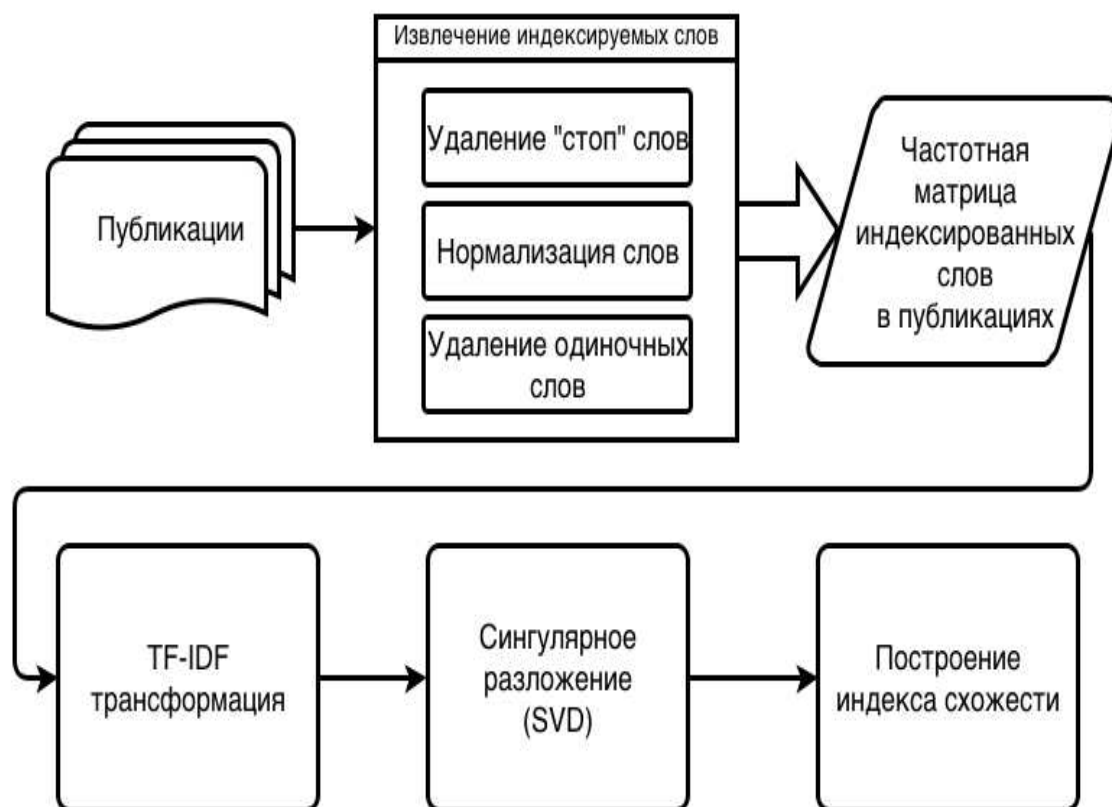


Рисунок 3.2 – Последовательность действий латентно-семантического анализа

Просмотрев список документов, можно заметить, что часть статей относится к медицинской тематике, а часть – к управлению проектами. Чтобы провести это разделение применяется латентно-семантический анализ.

Изначально имеется список тем, который нужно проанализировать и обработать с целью выделения индексируемых слов.

Таблица 3.1 – Список документов для примера работы латентно-семантического анализа

Д1	Когнитивные модели слабо структурированных проектов создания программных продуктов
Д2	Лекарственно индуцированные поражения печени: особенности выявления, постановки диагноза и ведения пациентов
Д3	Трансформация когнитивных карт в модели марковских процессов для проектов создания программного обеспечения
Д4	Особенности поражения печени при ВИЧ инфекции
Д5	Матричная диаграмма и сильная связность индикаторов ценности в проектах
Д6	Патогенетические механизмы прогрессирования сочетанных вирусных и

	алкогольных поражений печени
Д7	Разработка марковских моделей изменений состояния пациентов в проектах предоставления медицинских услуг
Д8	Решенные и нерешенные вопросы терапии неалкогольной жировой болезни печени в рамках метаболического синдрома
Д9	Анализ структурной модели компетенций по управлению проектами национального стандарта Украины

Анализ включает в себя:

1. удаление, так называемых, “стоп” слов, то есть, не имеющих смысловой нагрузки (предлоги, союзы и т.д.);
2. приведение слов к нормальному виду или стемминг-процесс нахождения основы слова (используется алгоритм Портера [51], который позволяет быстро определить основу слова);
3. удаление слов, встречающихся только один раз. Этот пункт не обязателен, но позволяет экономить ресурсы при расчетах.

На основе полученных индексируемых слов строится частотная матрица использования этих слов (табл. 3.2).

Таблица 3.2 – Частотная матрица использования индексируемых слов в документах

Индексируемые слова	Документы								
	Д1	Д2	Д3	Д4	Д5	Д6	Д7	Д8	Д9
когнитивн	1		1						
марковск			1				1		
модел	1		1				1		
особен		1		1					
пациент		1					1		
печен				1		1		1	
поражен		1		1		1			
программн	1		1						
проект	1		1		1		1		1
создан	1		1						

Для повышения качества анализа, следующий этап – трансформация матрицы с помощью модели TF-IDF (от англ. TF — term frequency, IDF — inverse document frequency) — статистическая мера, используемая для оценки важности слова в контексте документа, являющегося частью коллекции документов. Вес некоторого слова пропорционален количеству употребления этого слова в документе, и обратно пропорционален частоте употребления слова в других документах коллекции.

Следующий шаг, является основой латентно-семантического анализа – это сингулярное разложение полученной матрицы и построение индекса схожести, который вычисляется по расстоянию между индексируемыми словами и документами в k -мерном пространстве (табл. 3.3).

Таблица 3.3 – Результат сингулярного разложения частотной матрицы на ортогональные (U , V^t) и диагональную (S) матрицы

U									S									V^t									
-0.35	0.05	-0.32	0.25	-0.01	0.04	0.00	0.21	0.82	3.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.58	-0.05	-0.67	-0.01	-0.16	-0.01	-0.41	0.00	-0.16
-0.30	-0.02	0.29	-0.10	0.56	-0.64	0.23	0.20	0.00	0.00	2.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	-0.55	0.06	-0.65	0.00	-0.47	-0.10	-0.21	0.00
-0.47	0.01	0.08	0.05	0.34	0.26	-0.14	-0.76	0.00	0.00	0.00	1.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.36	0.37	-0.20	-0.21	0.14	-0.26	0.70	-0.23	0.14
-0.02	-0.47	0.09	0.38	-0.27	-0.47	-0.55	-0.14	0.00	0.00	0.00	0.00	1.35	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.55	0.13	-0.03	-0.38	-0.32	-0.26	-0.42	-0.38
-0.13	-0.26	0.61	0.21	0.15	0.52	-0.16	0.43	0.00	0.00	0.00	0.00	0.00	1.13	0.00	0.00	0.00	0.00	0.00	-0.25	-0.23	0.24	-0.08	-0.53	0.16	0.39	0.27	-0.53
-0.01	-0.52	-0.40	-0.57	0.31	0.13	-0.32	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.60	0.00	0.00	0.00	0.00	0.54	0.18	-0.52	-0.46	-0.10	0.32	0.13	0.22	-0.10
-0.02	-0.66	-0.06	0.15	-0.13	0.06	0.70	-0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.55	0.00	0.00	0.00	-0.21	-0.01	0.21	-0.30	0.05	0.70	-0.08	-0.57	0.05
-0.35	0.05	-0.32	0.25	-0.01	0.04	0.00	0.21	-0.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.31	0.00	0.00	-0.29	0.43	0.35	-0.47	0.12	-0.01	-0.29	0.52	0.12
-0.55	0.01	0.24	-0.52	-0.60	-0.06	0.03	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.71	0.00	0.00	0.00	-0.71
-0.35	0.05	-0.32	0.25	-0.01	0.04	0.00	0.21	-0.41	3.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.58	-0.05	-0.67	-0.01	-0.16	-0.01	-0.41	0.00	-0.16

Таблица 3.4 содержит координаты индексируемых слов и документов в двумерном пространстве ($k=2$), которые получены из ортогональных матриц после сингулярного разложения.

Таблица 3.4 – Координаты слов и документов в двумерном пространстве

Индексируемое слово	X	Y	Документ	X	Y
когнитивн	-0.35	0.05	Д1	-0.58	0.06
марковск	-0.30	-0.02	Д2	-0.05	-0.55
модел	-0.47	0.01	Д3	-0.67	0.06
особен	-0.02	-0.47	Д4	-0.01	-0.65
пациент	-0.13	-0.26	Д5	-0.16	0.00
печен	-0.01	-0.52	Д6	-0.01	-0.47
поражен	-0.02	-0.66	Д7	-0.41	-0.10
программн	-0.35	0.05	Д8	0.00	-0.21
проект	-0.55	0.01	Д9	-0.16	0.00
создан	-0.35	0.05			

На рис. 3.3 показано графическое представление индексируемых слов и документов. На рисунке хорошо видно, что часть документов принадлежит к одной тематике, а остальные – к другой. Проанализировав индексируемые слова, находящиеся рядом с документами, можно сделать вывод, что тематика 1 относится к управлению проектами (проект, модел, программн), а тематика 2 – к медицине (пациент, печен, поражен).

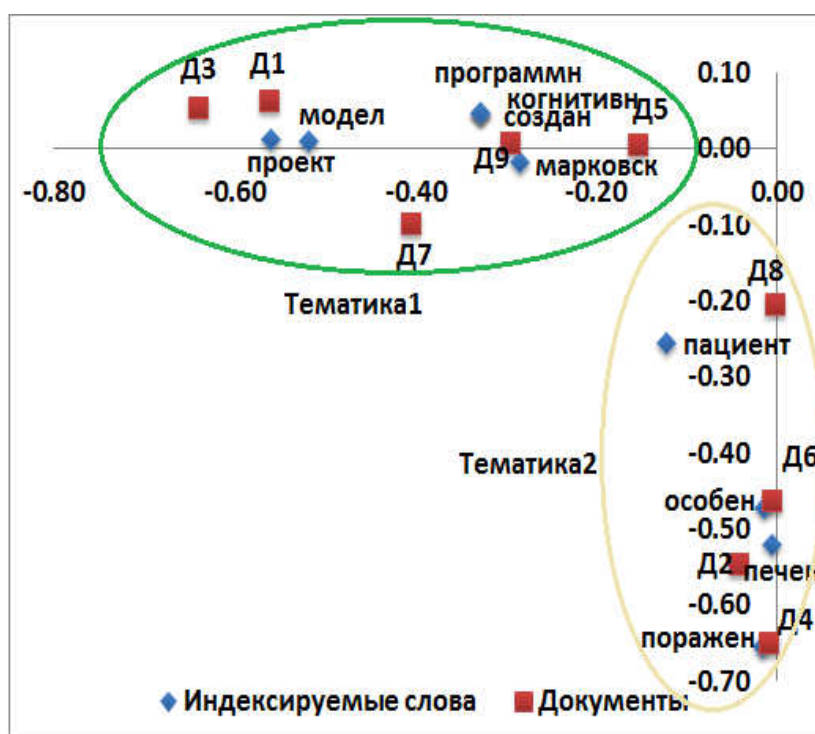


Рисунок 3.3 – Графическое представление распределения индексируемых слов и документов в двумерном пространстве

Следует подчеркнуть, что разрешение указанных выше задач должно строиться на основе законов проектного управления [34]. При этом необходимо учитывать общие подходы к организации терминологических систем научного знания [52], модели коммуникационных процессов [53], а также особенности организации компьютерных сетей [54].

Латентно-семантический анализ предоставляет довольно неплохие результаты сравнения различных документов по смыслу и дает возможность автоматической категоризации их. Будучи основанным на математических и статистических расчетах, этот подход является независимым от языка документов.

Применение латентно-семантического анализа в проекте извлечения информации из наукометрических баз данных позволяет решать проблему однофамильцев и выявить дубликаты.

3.3 Достоверность идентификации авторства научных публикаций на основе латентно-семантического анализа

Одним из этапов извлечения и сбора информации является ее обработка. В отличие от технического процесса извлечения, обработка может представлять собой интеллектуальную и даже творческую работу. Основной задачей на этом этапе является определение достоверности результатов [78]. Примером может служить следующая задача: заданы фамилия, имя и отчество (ФИО) автора и список публикаций, извлеченных по этим атрибутам; как определить статьи, только этого автора (так как атрибут запроса ФИО для разных авторов может совпадать).

Латентно-семантический анализ используется в обработке естественного языка, когнитивной науке и компьютерной лингвистике для решения подобных задач. Наиболее известными практическими реализациями ЛСА на сегодняшний день являются:

- SenseClusters. Основная функция – кластеризация схожих контекстов. Применяется при разрешении неоднозначности слов (в частности, имен), классификации документов разного рода (электронных писем, новостных статей), классификация лексики (нахождение синонимов, антонимов и других классов отношений) [65].

- S-Space. Основная функция – универсальное средство для построения и обработки векторной модели. Содержит реализации большого количества алгоритмов (разные векторные модели, некоторые методы их последующей обработки). Ориентировано на скорость работы, интуитивно понятное представление данных [66].

- Gensim. Наиболее надежное и эффективное программное обеспечение, которое реализует семантическое моделирование для обычного текста.

Предназначено специально для обработки больших коллекций документов, с использованием эффективных алгоритмов [67].

Проект по извлечению информации из наукометрических баз данных (НБД) подразумевает получение информации о публикациях, которые принадлежат конкретному автору, из наиболее известных НБД [11]. Выполнение поиска по заданному аргументу – ФИО – позволяет получить список публикаций, автором которых, по идее, является один человек. Но это не всегда верно, так как ФИО автора не может быть уникальным идентификатором записи. В мире могут существовать несколько авторов с одинаковыми ФИО. Добавим к этому тот факт, что чаще всего публикации содержат только инициалы с фамилией, поэтому вероятность нахождения публикаций нескольких авторов с идентичными ФИО, еще выше. Поэтому для выборки публикаций принадлежащих одному автору, нужно использовать дополнительную информацию из доступных полей структуры данных - название публикации. Название может отражать направление деятельности автора, а также это обязательное поле, которое не может быть пустым, в то время как остальные поля зачастую не доступны в тех или иных наукометрических базах [1]. На рис. 3.4 показан алгоритм использования ЛСА для классификации публикаций.



Рисунок 3.4 – Применение латентно-семантического анализа для определения авторства научных публикаций

Рассмотрим применение латентно-семантического анализа при анализе названий публикаций. В таблице 3.5 представлен пример результатов поиска публикаций для автора “Колесникова Е. В.”.

Таблица 3.5 – Фрагмент публикаций извлеченных по запросу
“Колесникова Е. В.”

№	НБД	Название публикации
1	Base-search	Лекарственно-индуцированные поражения печени: особенности выявления, постановки диагноза и ведения пациентов
2	Base-search	Современное состояние проблемы самоубийств в судебной медицине
3	Base-search	К вопросу о патоморфологических исследованиях нейроэндокринной системы при завершённых суицидах
4	Base-search	Теоретические исследования рабочего цикла гидравлического устройства ударного типа для ликвидации прихватов бурового снаряда в разведочных скважинах
5	Base-search	Гипоадипонектимия- ключевой фактор риска неалкогольной жировой болезни печени (обзор литературы)
6	Base-search	Особенности диагностики при подозрении на диффузную форму рака молочной железы
7*	Base-search	Трансформация когнитивных карт в модели марковских процессов для проектов создания программного обеспечения
8*	Base-search	Развитие теории проектного управления: обоснование закона К. В. Кошкина о завершении проектов

Как видно, часть статей связана с медицинской тематикой, но последние две публикации (отмеченные звездочками) относятся к совершенно иной предметной области. Если известно направление деятельности автора, то можно определить с некоторой погрешностью, какие из публикаций принадлежат данному автору. Для того, чтобы этот процесс автоматизировать, можно выделить ключевые слова из сферы предметной области автора и с помощью программы отобрать подходящие варианты. Но тут возникает проблема: нам нужен набор из множества слов, которые могут встречаться в названиях статей. Этот набор может быть слишком объёмным, что скажется на производительности. Кроме этого следует принимать во внимание, что

некоторые слова могут употребляться в разных контекстах с разным смыслом (проблема полисемии).

Чтобы решить эти проблемы используется латентно-семантический анализ, который позволяет выделить связь между ключевыми словами и набором документов (названий публикаций, в нашем случае). Допустим, задано ключевое слово “информация”. Применение латентно-семантического анализа позволяет установить скрытые связи, например, слова “программа” или “компьютер” близки к предметной области этого слова. Поэтому ЛСА позволяет получить не только список публикаций, где встречается слово “информация”, но и без этого слова с наиболее близкими по смыслу [50]. В табл. 3.6 представлен результат анализа по таким ключевым словам: “марковский”, “проект”, “информация”. Левая колонка отображает уровень схожести названия публикации с ключевыми словами. Схожесть определяется расстоянием между ключевым словом и документом в пространстве, построенном с помощью ЛСА (сингулярное разложение матрицы). Расчёт схожести документа и запроса рассчитывается по формуле косинуса:

$$\cos \theta = \frac{d_j \cdot q}{\|d_j\| \|q\|}, \quad (3.3)$$

где d_j – вектор j -ого документа в пространстве, q – вектор запроса в пространстве, $\|d_j\|$ и $\|q\|$ – норма векторов документа и запроса.

Таблица 3.6 – Результат латентно-семантического анализа по заданным ключевым словам

<i>%</i>	<i>Наукометрическая база</i>	<i>Название публикации</i>
88.75	Google academy	Матричная диаграмма и «сильная связность» индикаторов ценности в проектах
80.40	Google academy	Сетевые процессы управления проектами в контексте отображения состояний проекта
76.52	Base-search	Разработка марковских моделей изменений состояния пациентов в проектах предоставления медицинских

		услуг
69.47	Base-search	Трансформация когнитивных карт в модели марковских процессов для проектов создания программного обеспечения
55.57	Google academy	Управление знаниями в IT-проектах
51.27	Base-search	Составляющие поведенческой компетенции участника команды проекта на основе компетентностного подхода
47.99	Base-search	Анализ структурной модели компетенций по управлению проектами национального стандарта Украины
...		
-0.37	Base-search	К вопросу о патоморфологических исследованиях нейроэндокринной системы при завершённых суицидах

По этим результатам можно предположить, что публикации с высоким уровнем схожести принадлежат нашему автору, а с низким – иным авторам.

Еще одним вариантом применения латентно-семантического анализа является разбиение документов на некоторые группы, связанные по смыслу. Например, мы не можем выделить ключевые слова для направления научной деятельности. С помощью латентно-семантического анализа можно проанализировать список названий публикаций и разбить их, допустим, на 3 части и предоставить для каждой группы набор ключевых слов (табл. 3.7).

Таблица 3.7 – Ключевые слова, соответствующие смысловым группам

<i>№ группы</i>	<i>Ключевые слова</i>
1	печен, неалкогольн, жиров, болезн
2	систем, исследован, суицид, патоморфологическ
3	проект, процесс, управлен, состоян

По этому списку ключевых слов можно установить, что третья группа относится к тематике нашего автора. Выполнив поиск по ключевым словам этой группы, можно получить основную часть статей нашего автора, без публикаций медицинской тематики. Точнее, более высокий уровень схожести

будет получен для публикаций нашего автора [40]. Можно отбрасывать публикации, уровень схожести которых не превышает заранее заданный порог.

Ключевые слова, предложенные латентно-семантическим анализом, можно сохранить и в следующий раз использовать их при новом поиске публикаций этого автора. Таким образом, можно создать обучаемую систему в полуручном режиме и использовать в автоматическом.

Результат латентно-семантического анализа, конечно же, может иметь погрешность. Это хорошо заметно, когда ключевые слова можно отнести к различным предметным областям научной деятельности. Например, слово “проект” может использоваться в любой сфере: учебный проект, медицинский проект, управление проектами и т.д. При этом в документах с малым количеством слов, ключевые слова могут иметь большую весомость.

Использование ЛСА позволяет в некоторой степени автоматизировать и облегчить анализ документов. С его помощью можно выделять схожие по смыслу документы, классифицировать документы, а также извлекать ключевые слова, принадлежащие к различным смысловым группам.

На практике это позволяет решить проблему определения достоверности авторства публикаций. Несмотря на то, что в проекте по извлечению публикаций из наукометрических баз данных документы имеют названия из нескольких слов, применение ЛСА позволяет получить положительные результаты.

3.4 Модель латентного размещения Дирихле

Модель латентного размещения Дирихле (latent Dirichlet allocation, LDA, рис. 3.5) – предполагает, 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. Темы всех слов в документе предполагаются независимыми. В LDA, как и в PLSA,

документ может соответствовать не одной теме. Но LDA задаёт модель порождения, как слов, так и документов, поэтому появляется дополнительная возможность оценивать вероятности документов вне текстовой коллекции с помощью алгоритма вариационного вывода и семплирования по Гиббсу. В отличие от PLSA, в LDA число параметров не увеличивается с ростом числа документов в коллекции. Многочисленные расширения модели LDA устраняют некоторые её ограничения и улучшают производительность для конкретных задач [73].

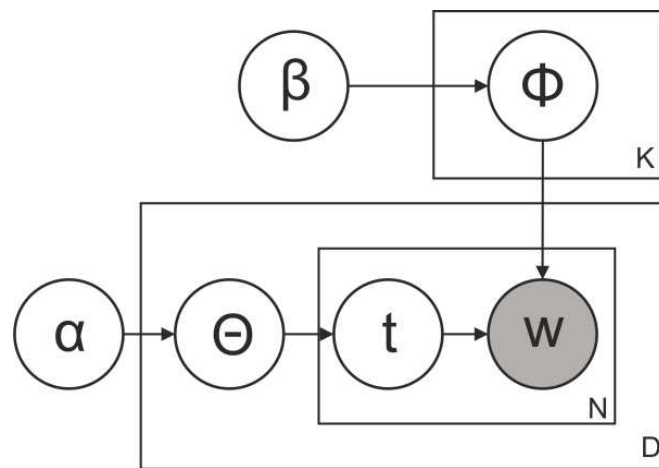


Рисунок 3.5 – Графическая вероятностная модель LDA, где w – слово (наблюдаемая переменная); t – тема (скрытая переменная); D – коллекция документов; N – длина документа в словах; K – количество тем в коллекции; θ – распределение тем в документе; ϕ – распределение слов в теме; α – априорное распределение Дирихле на параметры θ , β – априорное распределение Дирихле на параметры ϕ .

LDA безусловно лидирует среди вероятностных тематических моделей благодаря многочисленным обобщениям и приложениям к анализу коллекций текстовых документов:

- Автор-тематическая модель (author-topic model) представляет собой расширение LDA для совместного описания документов и авторов;
- Скрытая тематическая модель гипертекста (latent topic hypertext model, LTHM) описывает закон порождения ссылок в корпусе гипертекстов;

– Композитная модель HMM-LDA строится совместное описание синтаксиса и семантики текста. Скрытая марковская модель (HMM) описывает локальные закономерности между соседними словами, тогда как модель LDA даёт глобальное тематическое описание документа в целом [83].

Исходя из анализа существующих вероятностных тематических моделей, модель латентного размещения Дирихле (LDA) является хорошим кандидатом на использование в проекте по извлечению публикаций из наукометрических баз данных. После извлечения публикаций из наукометрических баз данных мы получаем список их названий. Задачей LDA является автоматическое определение тем, которые содержат эти названия. Применение LDA даст нам модель, показанную на рис. 3.6

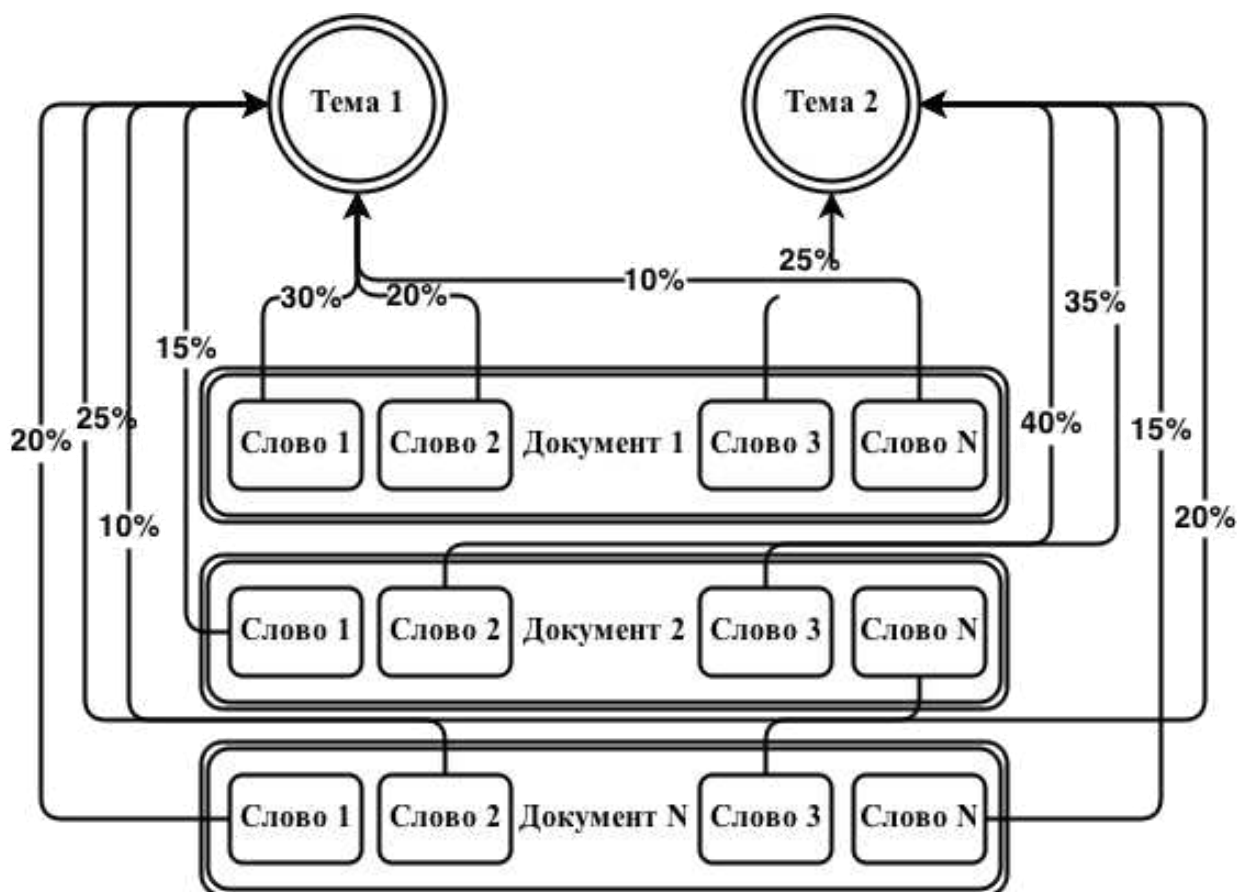


Рисунок 3.6 – Модель LDA с выделением двух тем.

На рисунке показаны распределения слов в двух темах. Исходя из этого, можно судить, что документ 1 больше относится к первой теме, чем ко второй

(30% +20% + 10% против 25%), а также получить список наиболее подходящих к теме слов.

Для идентификации параметров модели LDA по коллекции документов можно применить семплирование по Гиббсу – алгоритм для генерации выборки совместного распределения множества случайных величин. Он используется для оценки совместного распределения и для вычисления интегралов методом Монте-Карло. Допустим, мы хотим определить K тем в наборе документов, тогда алгоритм семплирования по Гиббсу можно описать так:

1. Для каждого слова из каждого документа присвоить случайным образом одну тему (t) из K возможных;
2. Для каждого слова из каждого документа вычислить:
 - $p(t|d)$ – пропорция слов в документе d , которые присвоены теме t ;
 - $p(w|t)$ – пропорция слова w во всех документах, присвоенного к теме t ;
 - присвоить слову w новую тему t с вероятностью $p(t|d) \cdot p(w|t)$.
3. Повторить второй пункт несколько раз (количество итераций также является входным параметром).

Метод LDA основан на следующей вероятностной модели:

$$p(d, w) = \sum_{t \in T} p(d) \cdot p(w | t) \cdot p(t | d), \quad (3.4)$$

где d – документ, t – тема, w – слово, T – множество тем, $p(d)$ – априорное распределение на множестве документов, $p(w|t)$ – условное распределение слова w в теме t , $p(t|d)$ – условное распределение темы t в документе d .

Для демонстрации результатов анализа, возьмем набор публикаций, полученный по параметру поиска «Яковенко В. Д.» и выполним сравнение с такими ключевыми словами: «система», «автоматизований». Процедура выполнена также с использованием LSA двумя повторениями. Результаты показаны в таблице 3.8.

Таблица 3.8 – Результат сравнения публикаций с ключевыми словами

LSA, %		LDA, %		Публикация
1	2	1	2	
82	82	100	100	Прогнозирование состояния системы управления качеством деятельности учебного заведения
0	0	52	53	К вопросу о причинно-следственных взаимосвязях в патогенезе хронического тонзиллита, как инфекционно-аллергического процесса
0	0	52	53	Некоторые закономерности соотношения дефицита барьерной функции миндалин и системного иммунитета при хроническом тонзиллите
99	99	41	54	Прогнозування стану системи керування якістю навчального закладу
87	87	96	36	Комп'ютерна реалізація системи автоматизованого управління навчальним процесом
88	88	96	45	Формалізація вимог до системи автоматизованого управління навчальним закладом

Из таблицы можно увидеть, что результаты LSA и LDA в некотором роде схожи, но из-за того, что LDA использует случайные величины, результаты могут различаться на одних и тех же входных документах. Также, из-за малого количества документов, LDA показывает достаточно большой процент схожести для документов, не подходящих к заданным ключевым словам.

Таким образом, можно сделать вывод, что для проекта по извлечению публикаций из наукометрических баз данных, латентно-семантический анализ подходит лучше, нежели вероятностная модель. Из-за небольшого объема, как публикаций, так и их содержимого (название в нашем случае), вероятностная модель латентного размещения Дирихле показывает худшие результаты. Учитывая, что одним из недостатков LSA является снижение скорости вычисления при увеличении объема данных, для этого проекта им можно пренебречь.

Латентное размещение Дирихле является базовой вероятностной тематической моделью и лидирует среди других благодаря многочисленным

обобщениям и приложениям к анализу коллекций текстовых документов. Базовые вероятностные тематические модели позволяют выявлять скрытую тематику документов на основе модели документа как мешка слов. В них также предполагается существование скрытых взаимосвязей между различными объектами, которые могут проявляться в структуре словоупотребления. Семантическая близость различных объектов может оцениваться путём сравнения их тематических векторов. Но, применив латентное размещение Дирихле к проекту по извлечению публикаций из наукометрических баз данных, замечено, что результаты хуже тех, которые получаются с использованием латентно-семантического анализа. Поэтому, не смотря на недостатки ЛСА, использование его в этом проекте оправдано.

3.5 Выводы

Модели со скрытыми (латентными) переменными являются особенно эффективными для выявления скрытых структур в текстовых коллекциях. Графические модели со скрытым тематическим слоем в последнее время успешно применялись в задачах поиска скрытых закономерностей в данных. Автоматическое выделение тематики текстов применялось для разбиения текстов по группам на основе семантической близости содержания. Одним из первых был предложен вероятностный латентный семантический анализ, вслед за ним – метод латентного размещения Дирихле и его многочисленные обобщения. Применение тематических моделей позволяет получить ответ на целый ряд нетривиальных вопросов, таких как: выявлять смысл или тематику документов по их содержанию, осуществлять классификацию документов на основе этих скрытых тематических закономерностей, выявлять научные интересы авторов и другие.

Латентно-семантический анализ представляет модель двухрежимного факторного анализа, которая основана на сингулярном разложении.

Сингулярное разложение представляет термины и документы в виде векторов в пространстве выбираемой размерности, а скалярное произведение между точками пространства – их схожесть. Латентно-семантический анализ хорошо справляется с проблемой синонимии, но частично с проблемой полисемии, потому, что каждое слово определяется одной точкой в пространстве.

Применение латентно-семантического анализа в данном исследовании позволяет решить проблемы идентификации публикаций определенного автора, а также выделить наиболее весомые ключевые слова, которые встречаются в названиях его публикаций. Названия публикаций содержат относительно малое количество слов, чтобы латентно-семантический анализ работал с минимально погрешностью. Не смотря на это, его применение все же показывает положительный результат.

Исследования раздела 3 опубликованы в работах автора [3-5, 12].

РАЗДЕЛ 4

РАЗРАБОТКА ПРОГРАММНОГО ПРОДУКТА ДЛЯ ИЗВЛЕЧЕНИЯ И ОБРАБОТКИ ИНФОРМАЦИИ ИЗ НАУКОМЕТРИЧЕСКИХ БАЗ ДАННЫХ

4.1 Требования и задачи программного продукта

В связи с требованиями к публикациям, которые относятся к диссертационной работе, возникает задача проверки наличия публикаций соискателя в международных наукометрических базах данных.

Задачей данного программного продукта является предоставить список публикаций соискателя, которые индексируются в международных наукометрических базах данных.

Одной из первых стадий разработки программного проекта является сбор информации, анализ, спецификация, и проверка требований к программному обеспечению. Программные требования – свойства программного обеспечения, которые должны быть надлежащим образом представлены в нём для решения конкретных практических задач. Опыт индустрии информационных технологий однозначно показывает, что вопросы, связанные с управлением требованиями, оказывают критически-важное влияние на программные проекты, в определенной степени – на сам факт возможности успешного завершения проектов. Требованиями к данному проекту являются:

- Извлечение информации из веб страниц;
- Критерием информации является ФИО автора;
- Работа с наиболее распространенными наукометрическими базами данных: Scopus, Web of Science;
- Обработка результатов с целью определения нерелевантной информации и фильтрации ее;
- Предоставление информации пользователю.

4.2 Программный проект

Процесс определения архитектуры, компонентов, интерфейсов и других характеристик системы или ее компонентов называется проектированием. Результат процесса проектирования – дизайн. Проектирование есть инженерная деятельность, в которой надлежащим образом анализируются требования для создания описания внутренней структуры ПО и являющейся основой для его конструирования. Программный дизайн (как результат деятельности по проектированию) должен описывать архитектуру программного обеспечения, то есть представлять декомпозицию программной системы в виде организованной структуры компонент и интерфейсов между компонентами. Важнейшей характеристикой готовности дизайна является тот уровень детализации компонентов, который позволяет заняться их конструированием. Проектирование программных систем можно рассматривать как деятельность, результат которой состоит из двух составных частей:

- Архитектурный или высокоуровневый дизайн – описание высокоуровневой структуры и организации компонентов системы;
- Детализированный дизайн – описывающий каждый компонент в том объеме, который необходим для конструирования.

Разделяют следующие виды дизайна:

- D-дизайн – декомпозиция структуры программного обеспечения в виде набора фрагментов или компонент;
- FP-дизайн – семейство архитектурных представлений, базирующихся на шаблонах;
- I-дизайн – создание высоко-уровневой концепции, видения того, что из себя будет представлять программная система; данный вид дизайна является результатом процесса анализа требований и их трансформации в подходы к реализации.

Проектирование программного обеспечения в понимании программной инженерии подразумевает D- и FP-дизайн. I-дизайн в большей степени относится к работе с программными требованиями.

4.3 Высокоуровневый дизайн

Декомпозиция структуры программного проекта в виде набора компонент представлена на рис 4.1. Проект представляет собой программный комплекс из нескольких приложений, взаимодействие которых предоставляет сервис поиска и извлечения публикаций указанного автора. Основными компонентами системы являются:

- веб приложение smd;
- веб сервис scrapyd;
- дополнительные компоненты, с которыми работают основные это:
- сервер реляционной бд mysql;
- сервер NoSQL БД MongoDB.

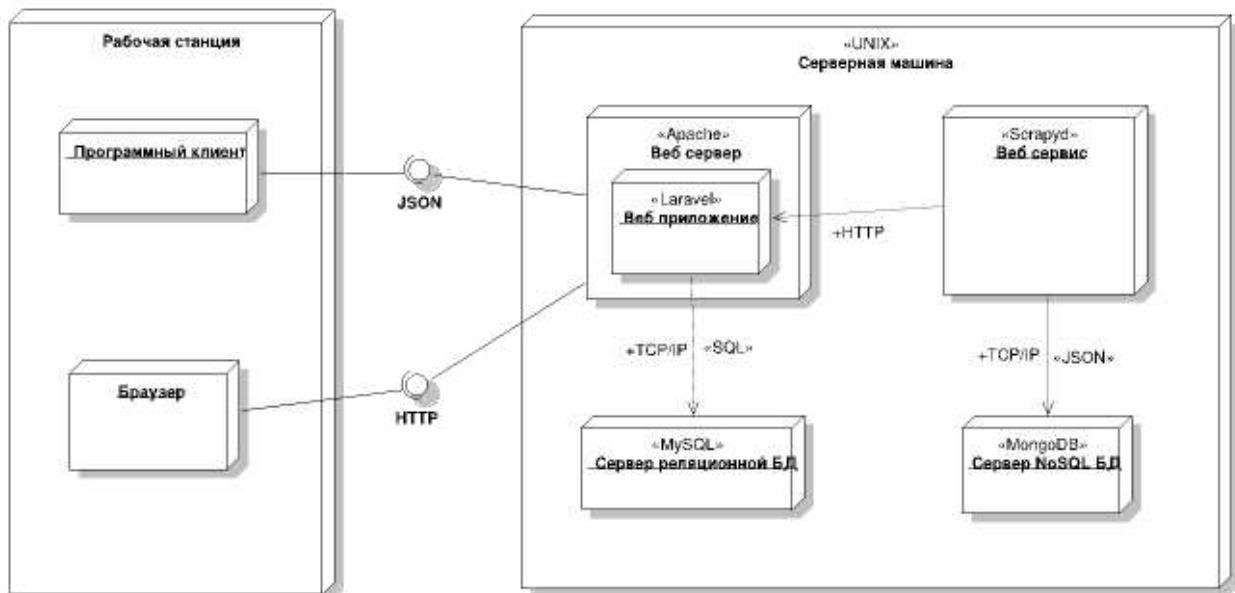


Рисунок 4.1 – Архитектура программного проекта по извлечению публикаций

Веб приложение SMD представляет собой графический интерфейс пользователя, а также предоставляет программный интерфейс для

использования поиска публикаций другими приложениями. Веб сервис Scrapyd представляет сервис по извлечению структурированных данных из НБД, а также управляет запуском соответствующих программ-пауков отдельной для каждой НБД. Таким образом, функционал программной системы разделен на отдельные модули — приложения, который работают независимо друг от друга. Веб приложение SMD использует сервис Scrapyd во время для поиска публикаций по запросу пользователя. Эти приложения общаются между собой по HTTP протоколу в JSON формате.

Веб приложение SMD использует реляционную базу данных (MySQL) в качестве хранилища данных, таких как информация о пользователях, список поддерживаемых НБД, история результатов поиска публикаций и др. Веб сервис Scrapyd использует документо-ориентированную базу данных (NoSQL) для временного хранения результатов поиска на внешнем диске, таким образом, не увеличивая объем использования оперативной памяти при извлечении большого количества публикаций. Доступ к базам данных предоставляют отдельные приложения — СУБД, с которыми приложения работают по протоколу TCP/IP.

Работа с приложением выполняется с помощью веб браузера. Также есть программный доступ к интерфейсу в формате JSON.

4.3.1 Основные варианты использования

На рис 4.2 отображены основные варианты использования приложения:

- регистрация пользователей в системе – создание учетной записи пользователя для возможности привязки найденных публикаций к пользователю.
- поиск публикаций – основной вариант использования, который включает в себя несколько этапов и участие двух актеров, которые инициируют действия. С одной стороны это пользователь, который запускает поиск по заданным

параметрам, с другой стороны это сервис поиска (scrapyd), который управляет этим процессом. Основные этапы поиска публикаций это извлечение информации, ее анализ (включая латентно-семантический) и сохранение результатов.

- история поиска публикаций – навигация по истории выполненных поисковых запросов.
- просмотр результатов поиска состоит из двух вариантов использования
- привязка публикаций к пользователю и отображение статистики по найденным публикациям или публикациям привязанных к пользователю.



Рисунок 4.2 – Варианты использования проекта по извлечению публикаций

Страница графического интерфейса результатов поиска публикаций включает в себя информацию о критериях поиска, количество найденных публикаций и идентифицированные публикации автора с помощью методов ранжирования, описанных в третьей главе. Поддерживается экспорт публикаций в систему Mendeley. Предусмотрено также отображение

нерелевантных результатов поиска, поскольку публикации автора могут относиться к различным областям знаний. Для повышения точности поиска публикаций автора поддерживается автоматическое формирование важных ключевых слов из названий избранных публикаций и сохранения их в профиле автора. Последующий поиск будет использовать эти ключевые слова для ранжирования результатов поиска публикаций.

Из известных аналогов данного программного продукта можно отметить программу Publish Or Perish, которая работает только с двумя источниками публикаций – Google Scholar и Microsoft Academic Search. Поэтому сравнение результатов поиска проведено с использованием одного источника – Google Scholar. Для оценки результатов поиска использованы следующие меры:

1. Точность – определяется как отношение числа релевантных документов, найденных системой, к общему числу найденных документов:

$$P = \frac{|D_{rel} \cap D_{retr}|}{|D_{retr}|}, \quad (4.1)$$

где D_{rel} – это множество релевантных документов в базе; D_{retr} – множество документов, найденных системой.

2. Полнота – отношение числа найденных релевантных документов, к общему числу релевантных документов в базе:

$$R = \frac{|D_{rel} \cap D_{retr}|}{|D_{rel}|}, \quad (4.2)$$

где D_{rel} – это множество релевантных документов в базе, а D_{retr} – множество документов, найденных системой.

3. Мера Ван Ризбергена, или F-мера – мера для совместной оценки точности и полноты, которая определяется как взвешенное гармоническое среднее точности P и полноты R :

$$F = \frac{(\beta^2 + 1)PR}{\beta^2 P + R}, \beta = \frac{(1 - \alpha)}{\alpha}, \alpha \in [0, 1], \beta \in [0, \infty] \quad (4.3)$$

При $\alpha = 1/2$ или $\beta = 1$ F-мера предоставляет одинаковый вес точности и полноте и называется сбалансированной степени, формула для нее упрощается:

$$F_1 = \frac{2PR}{P+R} \quad (4.4)$$

Число релевантных публикаций в базе Google Scholar тестового автора равняется 17. В табл. 4.1 приведены оценки результатов поиска для сравнения двух систем.

Таблица 4.1 – Сравнение результатов поиска публикаций

Система		Найдено релевантных публикаций/всего	Точность	Полнота	Мера Ван Ризбергена
Publish Or Perish		15/155	0.096	0.882	0.173
SMD	LSA	12/15	0.800	0.706	0.750
	LDA	9/10	0.900	0.529	0.666

На основе этих оценок можно сделать вывод о большей эффективности [79] разработанного программного комплекса по сравнению с программой Publish Or Perish. Значение сбалансированной F-меры SMD в несколько раз превышает показатель Publish Or Perish. Но следует отметить, что эта точность зависит от ключевых слов. Поэтому предусмотрено автоматическое извлечение наиболее весомых ключевых слов из названий публикаций. После первого поиска в системе, извлеченные ключевые слова сохраняются и при последующем поиске будут учтены, тем самым повышая точность идентификации публикаций автора.

4.3.2 Используемые средства и технологии

В данном проекте используются несколько языков программирования, различные библиотеки и приложения, которые отображены в таблице 4.2. Колонка «Где используется» показывает какой из двух основных компонентов использует ту или иную технологию.

Таблица 4.2 – Используемые средства и технологии

Категория	Значение	Где используется
Языки программирования	PHP, Javascript	Веб приложение
	Python	Веб сервис
Фреймворки, библиотеки	Laravel	Веб приложение
	Guzzle	Веб приложение
	Scrapy	Веб сервис
	Selenium WebDriver	Веб сервис
Внешние приложения	Gensim	Веб сервис
	NLTK	Веб сервис
	Apache web server	Веб приложение
	PhantomJS	Веб сервис
	MySQL server	Веб приложение
	MongoDB server	Веб сервис
	Scrapyd	Веб сервис

PHP (PHP: Hypertext Preprocessor) – скриптовый язык программирования общего назначения, интенсивно применяемый для разработки веб-приложений.

JavaScript – прототипно-ориентированный сценарный язык программирования. JavaScript обычно используется как встраиваемый язык для программного доступа к объектам приложений. Наиболее широкое применение находит в браузерах как язык сценариев для придания интерактивности веб-страницам.

Python – высокоуровневый язык программирования общего назначения, ориентированный на повышение производительности разработчика и читаемости кода.

Laravel – бесплатный веб-фреймворк с открытым кодом, предназначенный для разработки с использованием архитектурной модели MVC (Model View Controller — модель-представление-контроллер).

Guzzle – библиотека для PHP с помощью которой легко слать HTTP запросы и незатруднительно интегрировать приложение с веб сервисами.

Scrapy это фреймворк для обхода веб-сайтов и извлечения структурированных данных, которые могут быть использованы для широкого приложений.

Selenium – это инструмент для тестирования Web-приложений. **Selenium WebDriver API** используется для доступа к браузеру.

Gensim является библиотекой на языке программирование Python и предназначена для автоматического извлечения семантических тем из документов. Алгоритмы в gensim: латентного семантический анализ, латентный распределения Дирихле.

NLTK (Natural Language Toolkit) — набор библиотек и программ для символической и статистической обработки естественного языка на языке программирования Python.

Apache HTTP-сервер – свободный веб-сервер.

PhantomJS — скриптовый браузер без графического интерфейса, используемый для автоматизации взаимодействия с веб-страницами.

MySQL – свободная реляционная система управления базами данных.

MongoDB – документо-ориентированная система управления базами данных (СУБД) с открытым исходным кодом, не требующая описания схемы таблиц.

Scrapyd представляет собой приложение для развертывания и запуска SCRAPY пауков. Это позволяет развертывать ваши проекты и контролировать своих пауков с помощью JSON API.

4.4 Детализированный дизайн

4.4.1 Веб приложение SMD

4.4.1.1 Внутренняя структура

Веб приложение SMD имеет структуру Модель-Представление-Контроллер, где данные (модель), отображение их (представление) и

управление ими (контроллер) осуществляют отдельные компоненты. На рис. 4.3 показаны основные компоненты приложения. Компонент “роутер” отвечает за перенаправление запросов в соответствующий контроллер. Компонент “Миграции” позволяет создавать/изменять структуру данных, а наполнитель данных используется для генерации общих или тестовых данных. Данная структура приложения определена фреймворком Laravel, на основе которого оно построено.

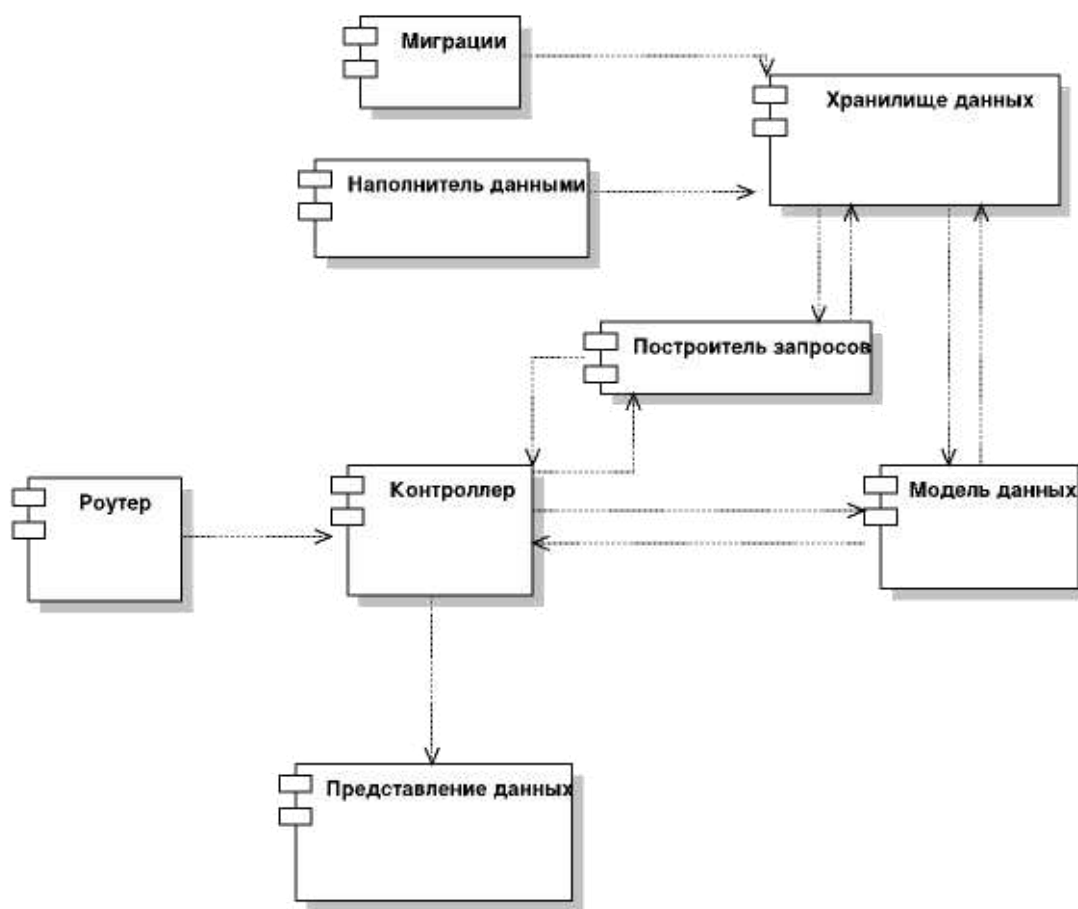


Рисунок 4.3 – Структура веб приложения SMD

4.4.1.2 Описание интерфейсов

В таблице 4.3 представлен веб интерфейс данного приложения. Для каждой страницы подробно описано возвращаемое значение и входные параметры.

Таблица 4.3 – Интерфейс веб приложения SMD

Метод	URL шаблон	Описание
GET	/	Возвращает начальную страницу
GET	/home	Возвращает домашнюю страницу вошедшего пользователя или перенаправляет на страницу входа
GET	/auth/login	Возвращает страницу для входа зарегистрированного пользователя
POST	/auth/login	Осуществляет вход зарегистрированного пользователя и перенаправляет на домашнюю страницу в случае успешного входа. Параметры запроса: <ul style="list-style-type: none"> – email – электронный адрес пользователя – password – пароль пользователя – remember – запомнить вошедшего пользователя
GET	/auth/register	Возвращает страницу регистрации пользователя
POST	/auth/register	Осуществляет регистрацию пользователя и перенаправляет на домашнюю страницу в случае успешной регистрации. Параметры запроса: <ul style="list-style-type: none"> – surname – фамилия пользователя – name – имя пользователя – farther_name – отчество пользователя – email – электронный адрес пользователя – password – пароль пользователя – password_confirmation – подтверждение пароля
GET	/search	Возвращает страницу с формой для заполнения параметров поиска публикаций.
POST	/search	Проверка входных параметров и запуск процедуры поиска публикаций. Возвращает страницу процесса поиска. Параметры запроса: <ul style="list-style-type: none"> – surname[] – список фамилий автора – name[] – список имен автора – farther_name[] – список отчеств автора – databases[] – список наукометрических баз данных – keywords – ключевые слова публикаций автора
GET	/search/query? {criterias, smds}	Запуск поиска публикаций по заданному списку наукометрических баз данных и списку ФИО автора. Возвращает результат в JSON формате: { "uid": "%ID%" }

		<p>%ID% - идентификатор поиска</p> <p>Параметры запроса:</p> <ul style="list-style-type: none"> - criterias – список ФИО автора разделенных запятой - smds – список наукометрических баз данных, разделенных запятой (base-search, copernicus, doaj, elibrary, mlibrary, sciencedirect, scopus, sgoogle, springer, worldcat)
GET	/search/status?{uid}	<p>Проверка статуса поиска. Возвращает результат в JSON формате:</p> <pre>{ "status": "%CODE%", "common_status": "%STATUS%", "spiders":</pre>
GET	/search/status?{uid}	<pre>[{"status": "%STATUS%", "items_count": %COUNT%, "name": "%SMDNAME%", "uid": "%ID%" }, ...]</pre> <p>%CODE% - oklerror %STATUS% - processldone %COUNT% - количество найденных публикаций %SMDNAME% - наукометрическая база данных %ID% - идентификатор поиска в данной наукометрической базе</p> <p>Параметры запроса: uid – идентификатор поиска</p>
GET	/search/result?{uid}	<p>Возвращает результаты для завершенного поиска по его идентификатору. Возвращает результат в JSON формате:</p> <pre>{ "status": "%CODE%", "results": [{"items": [%ITEMS%], "items_count": %COUNT%, "name": "%SMDNAME%" }, ...] }</pre> <p>%CODE% - oklerror %STATUS% - processldone</p>
GET	/search/cancel?{uid}	<p>%COUNT% - количество найденных публикаций %SMDNAME% - наукометрическая база данных %ITEMS% - список результатов со следующим набором возможных полей: spider, author, title, desc, date, source, url.</p> <p>Параметры запроса: -uid – идентификатор поиска</p> <p>Прерывает поиск по его идентификатору. Возвращает результат в JSON формате:</p>

		<pre>{"status": "ok"}</pre> Параметры запроса: – uid – идентификатор поиска
GET	/search/canceljob?{uid}	Прерывает поиск по одной из наукометрических баз данных по его идентификатору. Возвращает результат в JSON формате: <pre>{"status": "ok"}</pre> Параметры запроса: – uid – идентификатор поиска в конкретной наукометрической базе данных
GET	/results	Возвращает страницу со списком истории поиска
POST	/results	Сохранение результатов поиска Параметры запроса: – smds – список наукометрических баз данных – criterias – список ФИО автора – keywords – ключевые слова публикаций автора – results – результаты поиска в JSON формате
GET	/results/{id}	Возвращает страницу результатов поиска Параметры запроса: – id – идентификатор результатов поиска

4.4.1.3 Динамическое поведение

На рис 4.4 показана диаграмма последовательности выполнения поиска публикаций с использованием интерфейса веб приложения SMD. Пользователь инициирует поиск по указанным параметрам, далее приложение запрашивает запуск поиска у веб сервиса Scrapy и ждет его завершения, периодически опрашивая статус. Когда поиск завершен, история и результаты сохраняются в хранилище приложения SMD и выполняется процедура ранжирования результатов поиска с помощью методов основанных на латентно-семантическом анализе и латентном размещение Дирихле. В конечном результате пользователю возвращается веб страница с найденными и проанализированными публикациями, которые далее можно сохранить в файл, в систему Mendeley либо просто распечатать.

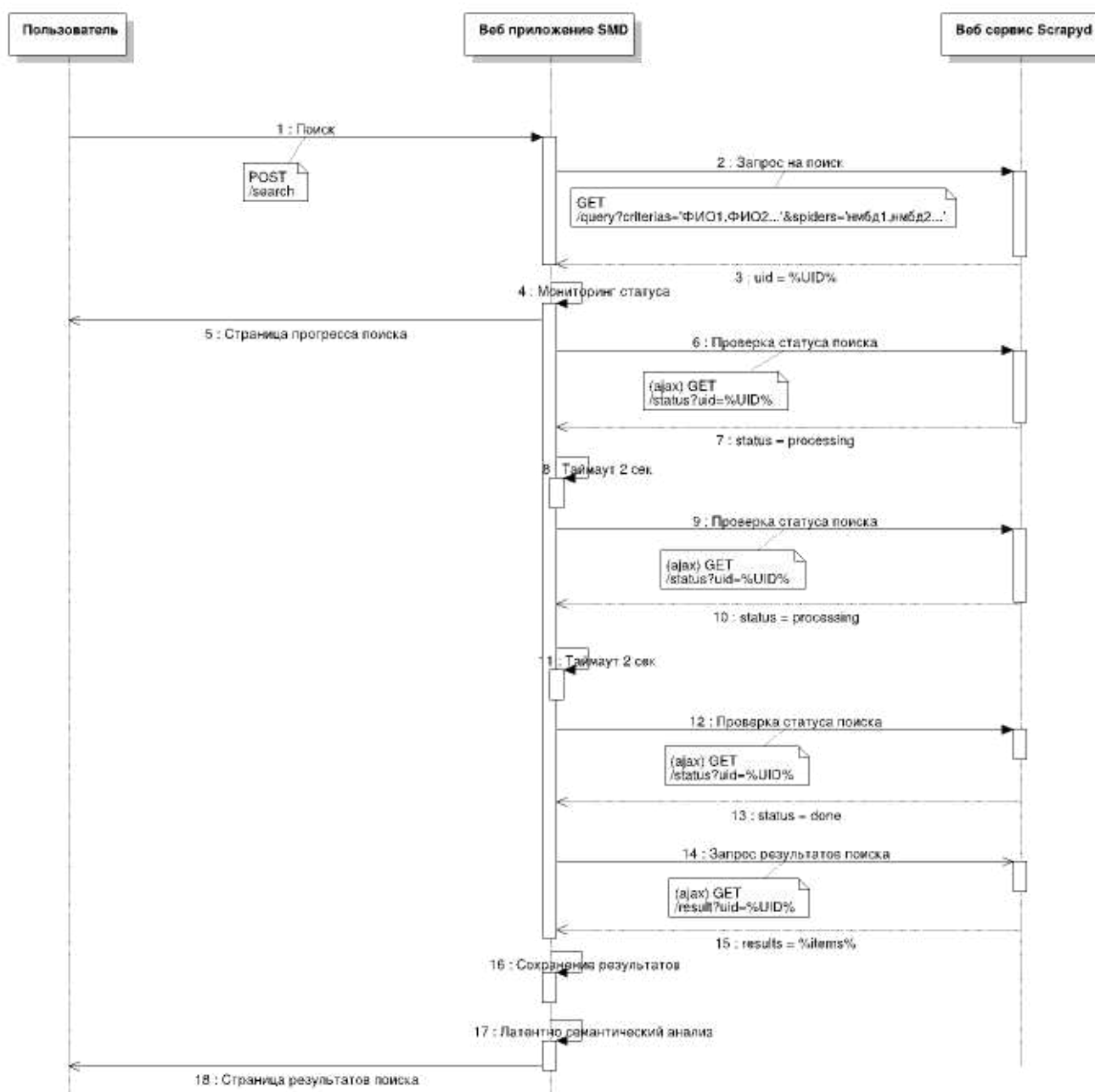


Рисунок 4.4 – Диаграмма последовательности выполнения поиска приложением SMD

4.4.1.4 Графический интерфейс пользователя

Интерфейс пользователя веб приложения представлен в виде веб страниц (html+css+javascript), которые браузер обрабатывает и отображает результат в графическом виде. На рис 4.5 показана главная или корневая страница приложения, где находится приветствие и список наукометрических баз данных

с которыми работает приложение. Панель навигации в шапке общая для всех страниц этого приложения, где можно переключиться на страницу задания параметров поиска или просмотреть историю предыдущих поисковых запросов с их результатами. Так же панель навигации содержит кнопки для входа зарегистрированных пользователей и регистрации новых.

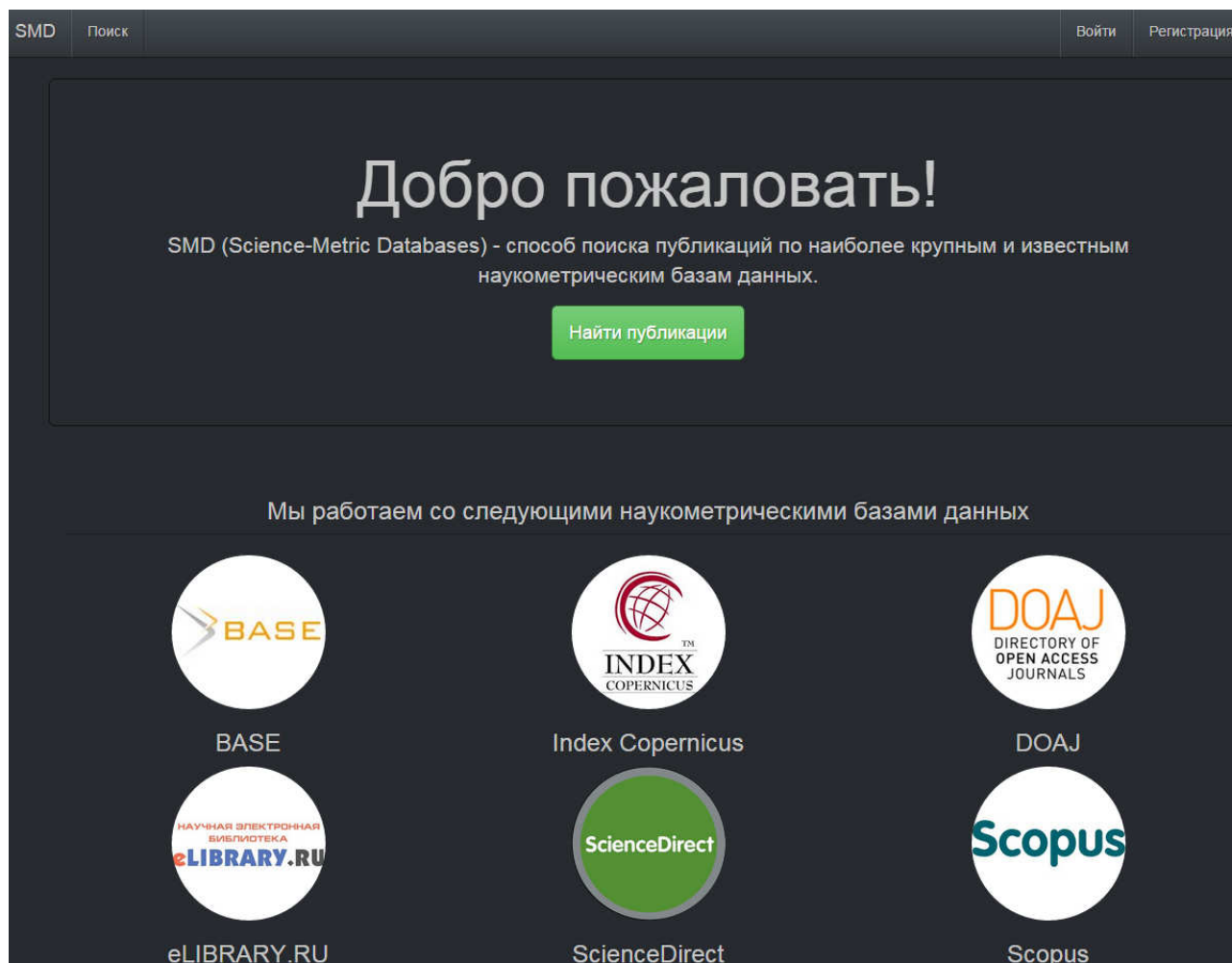


Рисунок 4.5 – Главная страница приложения

Страница формы поиска публикаций показана на рис 4.6. Здесь задается ФИО автора, ключевые слова его публикаций и выбирается список НБД, по которым будет выполняться поиск. ФИО автора можно задать несколькими вариантами, например, на разных языках. Для зарегистрированного пользователя автоматически подставляется информация из его учетной записи (варианты ФИО, ключевые слова).

Поиск публикаций

Коляда Андрей Сергеевич

+

баз, сист, веб, патентн, научн, scientometr, семантичesk, информaцион, информац, inform, публикац

BASE
 Index Copernicus
 DOAJ
 eLIBRARY.RU
 ScienceDirect

Scopus
 Google Scholar
 WorldCat
 Springer
 MLibrary

Выбрать все
 Очистить все

Рисунок 4.6 – Страница задания параметров поиска

На рис 4.7 показана страница истории поисковых запросов с количеством найденных результатов для каждого. По нажатии на пункт из списка выполнится переход на страницу результатов поиска, показанную на рис 4.8.

SMD	Поиск	История поиска	Коляда Андрей Сергеевич
2015-08-01 11:37:13	Автор(ы): Коляда Андрей Сергеевич, Kolyada Andrey Ключевые слова: баз, сист, веб, патентн, научн, scientometr, семантичesk, информaцион, информац, inform, публикац Базы: sgoogle		
2015-07-24 05:35:53	Автор(ы): Коляда Андрей Сергеевич, Kolyada Andrey Ключевые слова: баз, сист, веб, патентн, научн, scientometr, семантичesk, информaцион, информац, inform, публикац Базы: base-search; copernicus; doaj; elibrary; sciencedirect; scopus; sgoogle; worldcat; springer; mlibrary		
2015-07-24 05:29:51	Автор(ы): Коляда Андрей Сергеевич Ключевые слова: баз, сист, веб, патентн, научн, scientometr, семантичesk Базы: base-search; copernicus; doaj; elibrary; sciencedirect; scopus; sgoogle; worldcat; springer; mlibrary		

© ОНПУ, Кафедра УСБЖД 2015

Рисунок 4.7 – Страница истории поиска

Страница результатов содержит список публикаций с мета-информацией, а также статистику публикаций по НБД и по дате.

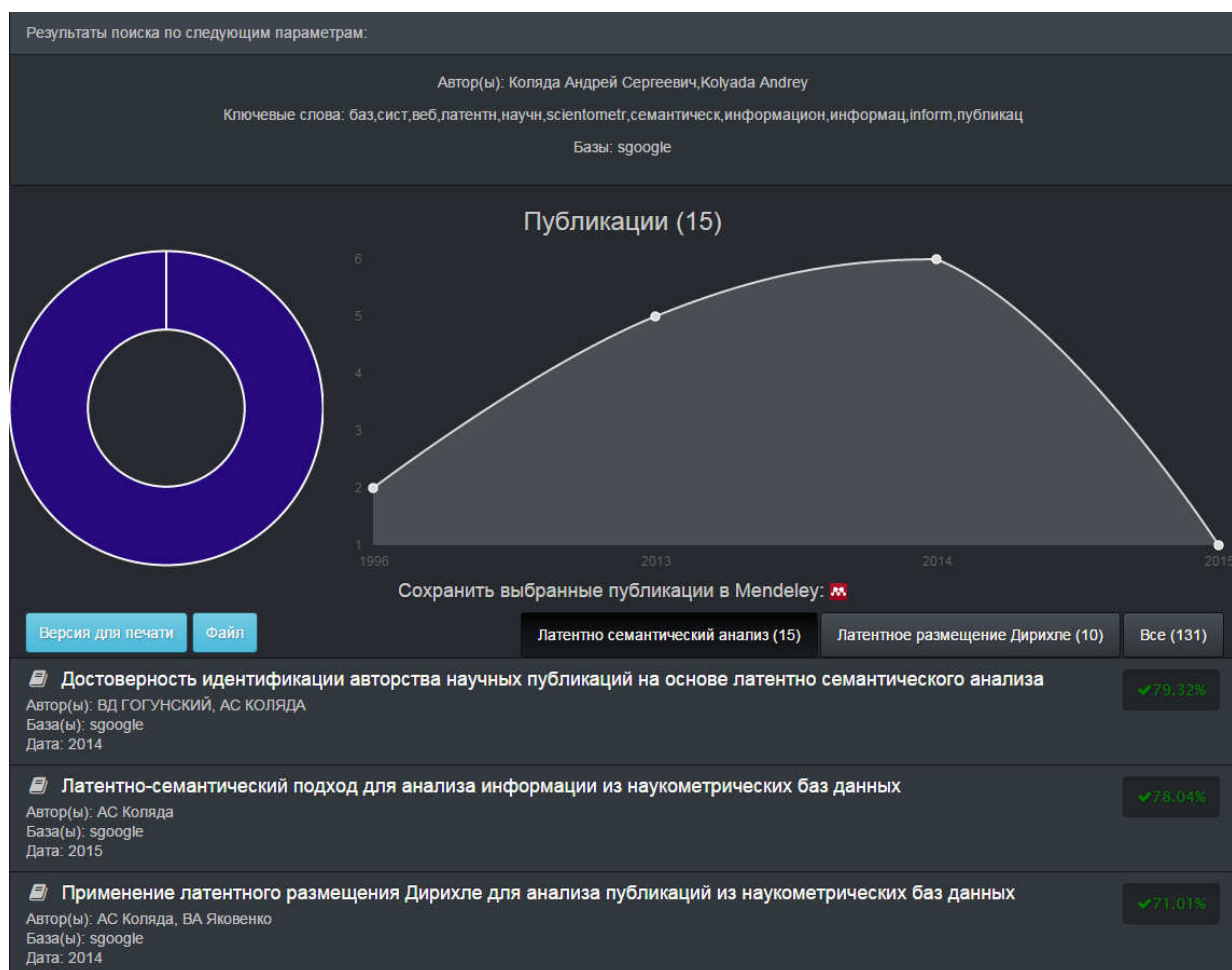


Рисунок 4.8 – Страница результатов поиска

4.4.2 Веб сервис Scrapyd

4.4.2.1 Внутренняя структура

Веб сервис Scrapyd (рис 4.9) предоставляет интерфейс запуска программ по извлечению данных для отдельных НБД, и, таким образом, скрывает сложную структуру этих программ, показанную на рис 4.10.

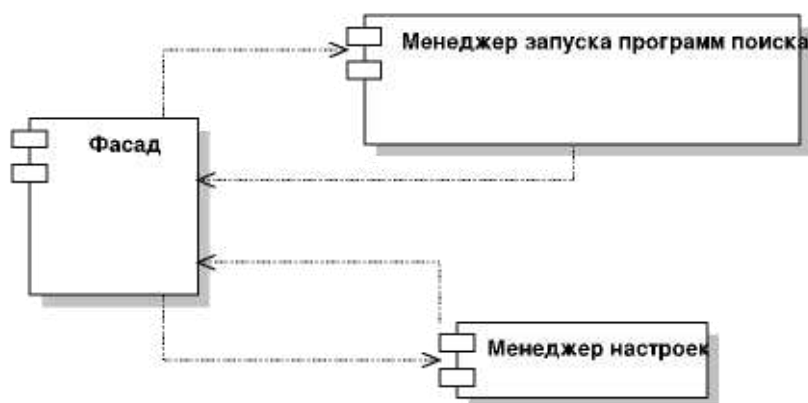


Рисунок 4.9 – Структура веб сервиса

Приложение по извлечению структурированных данных состоит из следующих компонентов:

- планировщик – управляет запуском отдельных пауков;
- загрузчик – отвечает за загрузку неструктурированных данных;
- пауки – программы, которые извлекают структурированные данные;
- обработчик результатов – анализ и преобразование извлеченных данных;
- контроллер – управляет всеми компонентами.

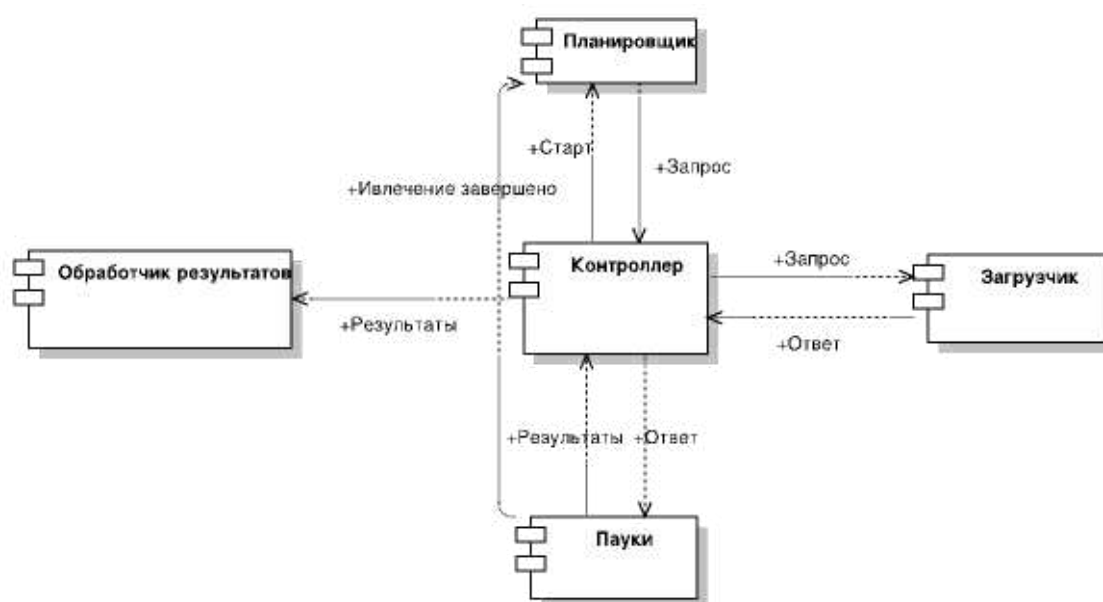


Рисунок 4.10 – Структура приложения извлечения метаданных публикаций

4.4.2.2 Описание интерфейсов

В таблице 4.4 отображен веб интерфейс веб сервиса с подробным описанием входных параметров и возвращаемых значений.

Таблица 4.4 – Интерфейс веб сервиса Scrapyd

Метод	URL шаблон	Описание
GET	/query?{criterias, spiders}	<p>Запуск поиска публикаций по заданному списку наукометрических баз данных и списку ФИО автора. Возвращает результат в JSON формате:</p> <pre>{ "status": "ok", "queryid": "%ID%" }</pre> <p>%ID% - идентификатор поиска</p> <p>Параметры запроса:</p> <ul style="list-style-type: none"> – criterias – список ФИО автора разделенных запятой – spiders – список наукометрических баз данных, разделенных запятой (base-search, copernicus, doaj, elibrary, mlibrary, sciencedirect, scopus, sgoogle, springer, worldcat)

GET	/status?{uid}	<p>Проверка статуса поиска. Возвращает результат в JSON формате:</p> <pre>{ "status": "%CODE%", "common_status": "%STATUS%", "spiders": [{ "status": "%STATUS%", "items_count": %COUNT%, "name": "%SMDNAME%", "uid": "%ID%" }, ...] }</pre> <p>%CODE% - oklerror %STATUS% - processldone %COUNT% - количество найденных публикаций %SMDNAME% - наукометрическая база данных %ID% - идентификатор поиска в данной наукометрической базе</p> <p>Параметры запроса:</p> <ul style="list-style-type: none"> – uid – идентификатор поиска
GET	/result?{uid}	<p>Возвращает результаты для завершенного поиска по его идентификатору. Возвращает результат в JSON формате:</p> <pre>{ "status": "%CODE%", "results": [</pre>
	/result?{uid}	<pre> { "items": [%ITEMS%], "items_count": %COUNT%, "name": "%SMDNAME%" }, ...] }</pre> <p>%CODE% - oklerror %STATUS% - processldone %COUNT% - количество найденных публикаций %SMDNAME% - наукометрическая база данных %ITEMS% - список результатов со следующим набором возможных полей: spider, author, title, desc, date, source, url.</p> <p>Параметры запроса:</p> <ul style="list-style-type: none"> – uid – идентификатор поиска
GET	/cancel?{uid}	<p>Прерывает поиск по его идентификатору. Возвращает результат в JSON формате:</p> <pre>{ "status": "ok" }</pre> <p>Параметры запроса:</p> <ul style="list-style-type: none"> – uid – идентификатор поиска
GET	/canceljob?{uid}	<p>Прерывает поиск по одной из наукометрических баз данных по его идентификатору. Возвращает результат в JSON формате:</p> <pre>{ "status": "ok" }</pre> <p>Параметры запроса:</p> <ul style="list-style-type: none"> – uid – идентификатор поиска в конкретной наукометрической базе данных

4.4.2.3 Динамическое поведение

На рис 4.11 показана диаграмма последовательности выполнения поиска публикаций веб сервисом Scaryd.

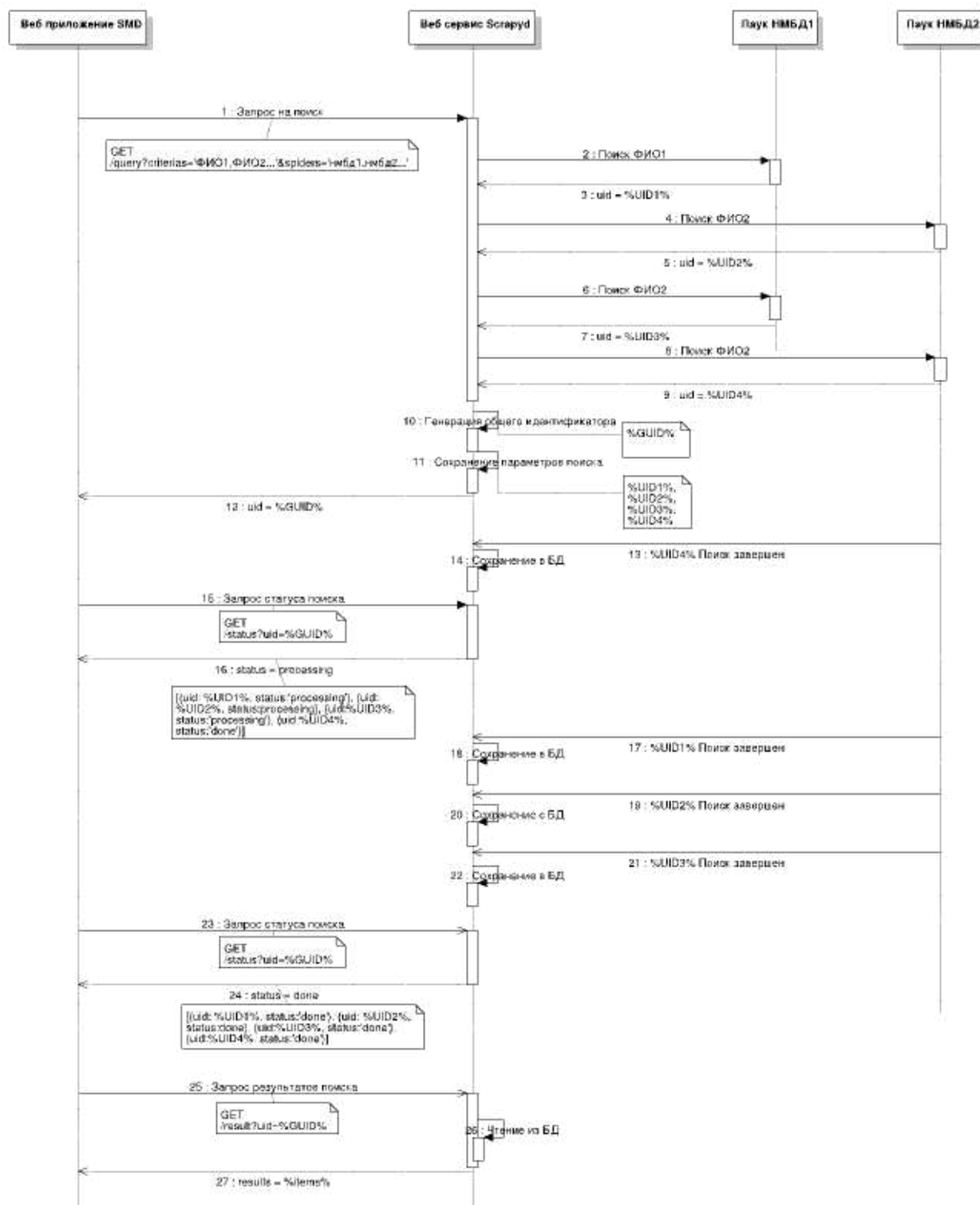


Рисунок 4.11 – Диаграмма последовательности выполнения поиска веб сервисом

4.4.3 Схема реляционной БД (MySQL)

Веб приложение использует реляционную базу данных для хранения метаданных публикаций, информации о пользователях, авторах, наукометрических базах с определением связей между ними. Схема этой базы данных показана на рис 4.12.

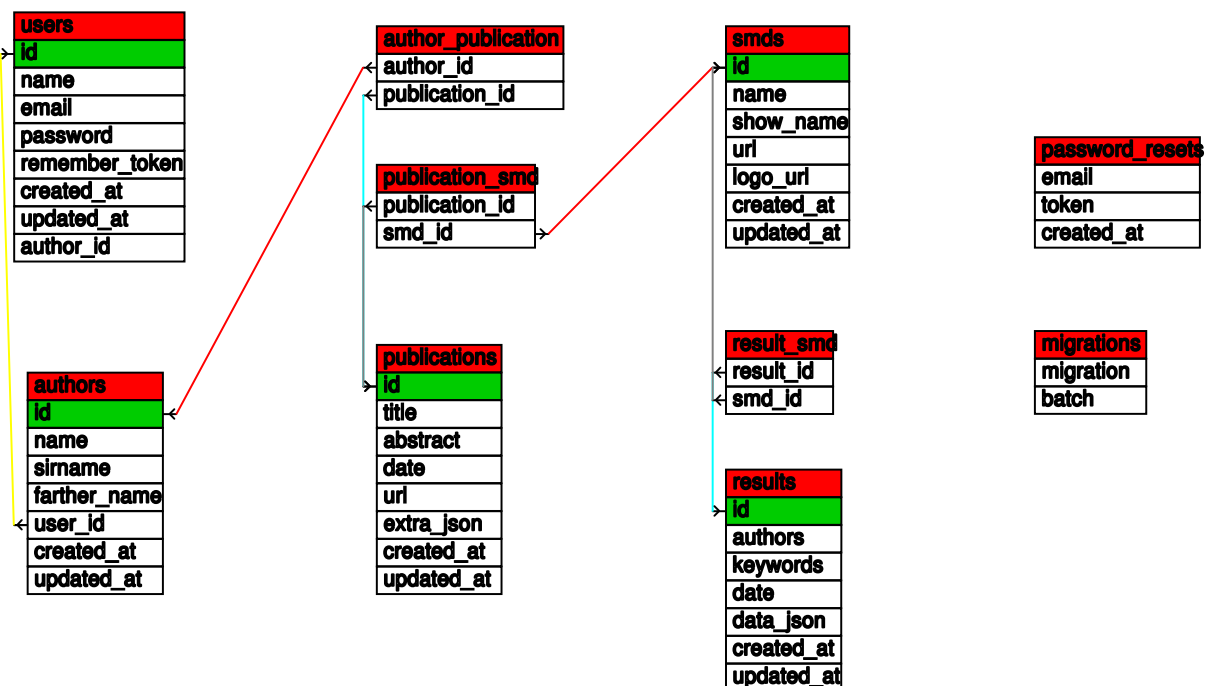


Рисунок 4.122 – Реляционная схема базы данных веб приложения SMD

Таблица `users` содержит информацию о зарегистрированных пользователях приложения (логин, пароль). Каждый зарегистрированный пользователь является автором и имеет связь с таблицей `authors` (один к одному).

Таблица `authors` содержит ФИО автора.

Таблица `publications` содержит метаданные научной публикации (название, дата, ссылка на источник, и др. поля).

У каждого автора может быть несколько публикаций. Эта связь, многие ко многим, представлена с помощью таблицы `author_publication`. Эта таблица содержит идентификатор автора и идентификатор научной публикации, которая относится к этому автору.

Таблица `smds` содержит список наукометрических баз данных, с которыми работает данное приложение.

Так как одна и та же публикация может индексироваться в разных наукометрических базах данных, существует таблица `publication_smd`, которая предоставляет связь многие ко многим для публикации и наукометрической базы данных, где она индексируется.

Таблица `results` содержит историю результатов поиска публикаций по заданному ФИО автора в выбранных наукометрических базах данных.

4.4.4 Схема NoSQL БД (MongoDB)

Веб сервис поиска публикаций использует NoSQL базу данных для временного хранения найденных извлеченных метаданных публикаций. Эта база данных состоит из двух коллекций:

- Запросы – состоит из списка документов “Запрос”
- Результаты – состоит из списка документов “Результат”

Документ “Запрос” содержит следующую структуру:

```
{
  "_id" : %OBJ_ID%,
  "spider_jobs" : [
    {
      "spider" : "%SMD_NAME%",
      "id" : "%JOB_ID%"
    },
    ...
  ],
  "criterias" : "%AUTHORS%",
  "uid" : "%SEARCH_ID%"
}
```

, где

`%OBJ_ID%` - идентификатор документа

`%SMD_NAME%` - имя наукометрической базы

`%JOB_ID%` - идентификатор поиска в данной наукометрической базе

`%AUTHORS%` - критерий поиска (список ФИО авторов)

`%SEARCH_ID%` - идентификатор данного поискового запроса

Документ “Результат” содержит следующую структуру:

```
{
  "_id" : %OBJ_ID%,
  "queryid" : "%SEARCH_ID%",
  "items" : [
    {
      "title" : "%PUB_TITLE%",
      "url" : "%PUB_URL%",
      "author" : [ %PUB_AUTHORS% ],
      "spider" : "%PUB_SMD_NAME%",
      "source" : "%PUB_SOURCE%",
      "date" : "%PUB_DATE%",
      "desc" : "%PUB_DESC%"
    },
    ...
  ]
}
```

, где

%OBJ_ID% - идентификатор документа

%SEARCH_ID% - идентификатор данного поискового запроса

%PUB_TITLE% - название публикации

%PUB_URL% - внешний URL на публикацию

%PUB_AUTHORS% - список авторов публикации

%PUB_SOURCE% - источник публикации

%PUB_DATE% - дата публикации

%PUB_DESC% - описание публикации

4.5 Применение и развертывание разработанного программного продукта

Данный программный продукт разработан как один из инструментов информационного обеспечения мониторинга публикационной активности ученых. Он предоставляет метаданные публикаций, которые индексируются в международных наукометрических базах данных. Основными особенностями программного продукта являются: возможность извлечения информации из

неструктурированных данных (Веб-страниц) и обработка этой информации с целью определения нерелевантной информации и фильтрации ее.

Использование данного программного обеспечения возможно в двух вариантах:

- индивидуальный поиск публикаций автора и сохранение их в свою коллекцию публикаций (например, Mendeley);
- мониторинг публикационной активности кафедр, университетов и других инстанций.

Второй вариант подразумевает использование данного программного продукта внешним приложением, с помощью программного интерфейса (API). Внешним приложением может быть информационная система с базой данных университетов, которой нужен способ извлечения метаданных публикаций из наукометрических баз данных для дальнейшей статистической обработки и мониторинга их. Программный интерфейс подробно описан в таблице 4.2 и состоит из следующих запросов:

- `/search/query?{criterias, smds}` – инициирует поиск по заданным авторам и наукометрическим базам данных и возвращает идентификатор данного поискового запроса;
- `/search/status?{uid}` – возвращает статус поиска по его идентификатору;
- `/search/result?{uid}` – возвращает список результатов по идентификатору поискового запроса;
- `/search/cancel?{uid}` – отменяет активный поиск по его идентификатору;

Таким образом, данный программный продукт можно использовать из веб-браузера для личных целей либо автоматизировано другими приложениями.

Система состоит из набора используемых программных компонентов и библиотек (табл. 4.1), что усложняет ее установку на отдельную рабочую станцию. Поэтому все компоненты установлены на отдельную виртуальную машину, которую можно развернуть на любом сервере, поддерживающем некоторый способ виртуализации. Веб или программный интерфейс,

описанный выше, позволяют пользоваться системой удаленно. Таким образом, на стороне пользователя отсутствует необходимость установки и настройки программного продукта.

Один экземпляр виртуальной машины с установленной и настроенной системой развернут на сервере кафедры УСБДЖ ОНПУ. Доступ по адресу <http://smd.opu.ua> (внутренний для ОНПУ IP-адрес: 10.53.29.1). Также, для обеспечения доступности системы в любое время и безотказной работы оборудования, развернут второй экземпляр виртуальной машины в «облаке». Доступ по адресу <http://smd.akolyada.tk>. Облачные вычисления – информационно-технологическая концепция, подразумевающая обеспечение повсеместного и удобного сетевого доступа по требованию к общему набору конфигурируемых вычислительных ресурсов (например, сетям передачи данных, серверам, устройствам хранения данных, приложениям и сервисам — как вместе, так и по отдельности), которые могут быть оперативно предоставлены и освобождены с минимальными эксплуатационными затратами или обращениями к провайдеру. Потребители облачных вычислений могут значительно уменьшить расходы на инфраструктуру информационных технологий (в краткосрочном и среднесрочном планах) и гибко реагировать на изменения вычислительных потребностей, используя свойства вычислительной эластичности облачных услуг. С точки зрения потребителя эти характеристики позволяют получить услуги с высоким уровнем доступности и низкими рисками неработоспособности, обеспечить быстрое масштабирование вычислительной системы благодаря эластичности без необходимости создания, обслуживания и модернизации собственной аппаратной инфраструктуры.

4.6 Выводы

Данный программный продукт разрабатывается как один из инструментов мониторинга публикационной активности научных сотрудников. Задачей его является предоставить метаданные публикаций соискателя, которые индексируются в международных наукометрических базах данных. Основными требованиями к программному продукту являются: возможность извлечения информации из неструктурированных данных (веб страниц) и обработка этой информации с целью определения нерелевантной информации и фильтрации ее.

Данный программный продукт состоит из веб приложения и веб сервиса, которые взаимодействуют между собой. Веб сервис занимается поиском и извлечением публикаций, а веб приложение предоставляет графический интерфейс пользователя, который отображает найденные публикации и содержит интерфейс инициализации поиска. Веб приложение также предоставляет программный интерфейс для возможного автоматизированного использования его.

Для реализации извлечения информации из веб страниц используется фреймворк Scrapy, который позволяет быстро создать программу паука. Для обработки динамических страниц используется безголовый браузер PhantomJS. Результаты извлеченных данных временно сохраняются в NoSQL базу данных MongoDB, а потом преобразуются в реляционные данные и хранятся в БД MySQL. За реализацию веб приложения отвечает веб фреймворк Laravel, с помощью которого за небольшое время можно создать стабильно работающее приложение.

Исследования раздела 4 опубликованы в работах автора [1–5, 11, 14, 15].

ОБЩИЕ ВЫВОДЫ И РЕКОМЕНДАЦИИ

В диссертации решена актуальная научно-прикладная задача теоретического обоснования моделей и методов анализа контента Веб страниц с имитацией работы пользователей для автоматизированного извлечения метаданных научных статей из наукометрических баз данных с помощью разработанных программных инструментов.

Получены следующие результаты:

1 Вклад в теоретические основы информационных технологий:

1.1 На основе анализа опубликованных работ и существующих программных продуктов установлено, что поиск публикаций в наукометрических базах данных, как правило, осуществляется только в пределах отдельных баз данных или репозиториев, что не позволяет определить интегральную оценку публикационной активности ученых;

1.2 Выполнена формализация информационной технологии для задач управления поиском метаданных публикаций в наукометрических базах данных, включая современную компьютерную систему накопления, переработки и хранения информации, что позволяет разработать и внедрить Интернет-технологии для построения сервис-ориентированной системы информационного обеспечения конечных пользователей;

1.3 Обоснована и разработана информационно-поисковая система автоматизации извлечения метаданных публикаций из распространенных наукометрических баз данных, включая программные инструменты извлечения и анализа контента веб страниц, что позволяет выполнить интегральную оценку публикационной активности авторов научных публикаций;

1.4 Усовершенствован метод Дирихле и модель латентно-семантического анализа, содержащих вероятностные оценки и инструментальные средства классификации и определения достоверности информации, которая извлекается из контента веб страниц и основаны на анализе скрытых переменных для

выявления связей в наборе названий публикаций, позволяющей достоверно идентифицировать публикации конкретных авторов.

2 Вклад в методы построения информационно-поисковых систем:

2.1 Предложена концепция построения информационно-поисковых систем и способов информационного обеспечения пользователей, основанной на информационной технологии извлечения и анализа контента веб страниц наукометрических баз данных, позволяет выполнять мониторинг интегральной публикационной активности, как отдельных ученых, так и научных коллективов.

2.2 Разработанные программные инструменты извлечения информации из веб страниц, которые конструируются динамически на стороне пользователя (клиента), что позволяет построить информационную технологию извлечения контента с элементами интеллектуальности в условиях неопределенности.

2.3 Разработан программный продукт, реализующий информационную технологию поиска публикаций ученых из наиболее известных наукометрических базах данных; программный продукт может быть полезным, как учебными заведениями, так и отдельным ученым, которым полезно узнать какие из их публикаций индексируются определенными наукометрическими базами данных.

3 Создание предпосылок для дальнейших исследований:

3.1 Результаты диссертационных исследований могут служить основой для развития информационных технологий по обеспечению информационных потребностей отдельных ученых с созданием информационно-поисковых систем для большего числа наукометрических баз данных.

3.2 Предложена и разработана информационная технология, в работе ориентирована на обеспечение личных информационных потребностей отдельных ученых, может быть формализована, как программный интерфейс (API), для включения в другие программные комплексы, например, для мониторинга публикационной активности лабораторий, кафедр, институтов.

ПЕРЕЧЕНЬ ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. **Коляда, А. С.** Автоматизація вилучення інформації із науко метричних баз даних [Текст] / А. С. Коляда, В. Д. Гогунський // Управління розвитком складних систем. – 2013. – № 16. – С. 96 – 99.
2. **Коляда, А. С.** Вилучення інформації із слабоструктурованих веб сторінок / А. С. Коляда, В. Д. Гогунський // Східно-Європейський журнал передових технологій. – № 1/9 (67). – Харків : Технолог. центр, 2014 – С. 51 – 54.
3. **Коляда, А. С.** Латентно-семантичний підхід для аналізу інформації із наукометрических баз даних / А. С. Коляда // Управління розвитком складних систем. – 2014. – Вып. 17. – С. 90 – 94.
4. **Коляда, А. С.** Достовірність ідентифікації авторства научних публікацій на основі латентно-семантичного аналізу / А. С. Коляда, В. Д. Гогунський // Східно-Європейський журнал передових технологій. – № 3/2 (69). – Харків : Технолог. центр, 2014 – С. 36 – 40.
5. **Коляда, А. С.** Застосування латентного розміщення Дирихле для аналізу публікацій із наукометричних баз даних / А. С. Коляда, В.А. Яковенко, В. Д. Гогунський // Праці Одес. політехн. ун-ту. 2014. – № 1 (43). – С. 186 – 191.
6. Гогунський В. Д. Наукометричні дані наукового видання «Управління розвитком складних систем» / В.Д. Гогунський, **А.С. Коляда**, В.А. Яковенко // Управління розвитком складних систем. – 2014. – № 19. – С. 6 – 11.
7. Gogunsky V.D. The development of the system concept of scientometric databases / V.D. Gogunsky, V.O. Iakovenko, **A.S. Kolyada** // Management of Development of Complex Systems, Number 20, 2014. – pp. 143 – 147.
8. Гогунський В.Д. . Особливості цитування наукових публікацій у інтернет-просторі / В.Д. Гогунський, В.О. Яковенко, А.С. Коляда // Наук.-метод. семінар: „Шляхи реалізації кредитно-модульної системи організації навчального процесу і тестових форм контролю знань студентів”. – Вип. 10. — Одеса : Наука і техніка, 2015. — С. 28 – 33.

9. Коляда А. С. Сучасні тенденції розвитку систем автоматизації технологічних процесів за наукометричними даними Scopus / А.С. Коляда, В.Д. Гогунський, В. А. Волобоев // Наук.-метод. семінар: „Шляхи реалізації кредитно-модульної системи організації навчального процесу і тестових форм контролю знань студентів”. – Вип. 10. — Одеса : Наука і техніка, 2015. — С. 46 – 52.

10. Гогунський, В. Д. Наукометричні бази: характеристика, можливості і завдання / В. Д. Гогунський, Г.О. Оборський, А. С. Коляда // Наук.-метод. семінар: „Шляхи реалізації кредитно-модульної системи організації навчального процесу і тестових форм контролю знань студентів”. – Вип. 8. — Одеса : Наука і техніка, 2014. — С. 3 — 12.

11. Коляда А. С. Розробка програмного проекту для вилучення і обробки інформації із наукометричних баз даних / А.С. Коляда., В. Д. Гогунський // Інформ. технології в освіті, науці та виробництві : зб. – Вип. 2 (7). – Одеса : АО Бахва, 2014 – С. 191 – 195.

12. Коляда А. С. Латентно-семантичний аналіз інформації із наукометричних баз. / А. С. Коляда // Наук.-метод. семінар: „Шляхи реалізації кредитно-модульної системи організації навчального процесу і тестових форм контролю знань студентів”. – Вип. 9. — Одеса : Наука і техніка, 2014. — С. 30 – 36.

13. Гогунський В. Д. Розробка наукометричних баз даних / В.Д. Гогунський, В.А. Яковенко, А.С. Коляда // Автоматизація: проблеми, ідеї, рішення: матеріалм міжнар. науч. - техн. конф. Севастополь, 8-12 вересня 2014 г. / Севастоп. нац. техн. ун-т; науч. ред. В.Я. Копп - Севастополь:, СевНТУ, 2014. – 184 с. – С. 111 — 113.

14. Гогунський В.Д. Проектування системи моніторингу публікацій науковців в наукометричних базах даних / В.Д. Гогунський, А.С. Коляда, В.А. Яковенко // «Молодь у світі сучасних технологій за тематикою: Теоретико-методологічні та наук. практи. засади управління проектами підвищення

конкурентоспроможності територій» // Матеріали III Міжнар. науково-практ. конф. студентів, аспірантів та молодих вчених (Херсон, 5-6 червня 2014 р.) / за заг. ред. Н.А. Соколова. Херсонський нац. техн. ун-т. – Херсон: ХНТУ, 2014. – (145 с.) С. 28–33.

15. Коляда, А. С. Розробка проекту інформаційно-аналітичної системи вилучення і обробки інформації із наукометричних баз даних / А. С. Коляда, А. А. Негрі, Е. В. Колеснікова // Управління проектами: стан та перспективи. Матеріали IX Міжнар. Наук.-практ. Конф. — Миколаїв : НУК, 2013. — С. 348.

16. Коляда А. С. Запобігання інформаційної небезпеки/ А. С. Коляда, В. Д. Гогунський // Наук.-метод. Семінар: „Шляхи реалізації кредитно-модульної системи організації навчального процесу і тестових форм контролю знань студентів”. – Вип. 7. — Одеса : Наука і техніка, 2013. — С. 54 — 57.

17. Коляда А.С. Розробка інтелектуальних систем навчання із застосуванням програмних агентів / А.С. Коляда, В.Д. Гогунський // «Шляхи реалізації кредитно-модульної системи ...». – Вип. 5. — Одеса: ОНПУ, 2011. – С. 45 – 49.

18. Коляда А. С. Ефективність використання адаптивних підходів при розробці програмного забезпечення / А.С. Коляда, С.Н. Ковешніков, В.Д. Гогунський // Інформ. Технології в освіті, науці та виробництві : зб. Наук. Праць. – Вип. 1. – Одеса : АО Бахва, 2012 –132 с. (С. 29 – 33)

19. Костирко, Т. Н. Університети України: приєднання до руху відкритого доступу [Текст] // Вісник ОНУ. – Том 16. – Випуск 1/2 (5/6). – 2011. – С. 283 – 289.

20. Рейтинг лучших университетов мира по версии QS [Электронный ресурс] — Режим доступа: <http://gtmarket.ru/ratings/qs-world-university-rankings/info>

21. Шанхайский рейтинг лучших вузов мира: ARWU [Электронный ресурс] — Режим доступа: <http://www.education-medelle.com/articles/schankhajskij-rejting-luchschikh-vuzov-mira-arwu.html>

22. Оборський, Г. О. Нові тенденції і завдання щодо підготовки науковців вищої кваліфікації [Текст] / Г. О. Оборський, В. Д. Гогунський // Інформаційні технології в освіті, науці та виробництві. — Вип. 2. — Одеса : АО Бахва, 2013. — С. 15 — 22.
23. Price, Derek J. de Solla (1963). Little science, big science. New York: Columbia University Press.
24. Garfield E (1955). "Citation indexes for science: A new dimension in documentation through association of ideas". Science 122: pp. 108–11.
25. Білощицький, А.О. Наукометричні бази та індикатори цитування наукових публікацій / А. О. Білощицький, В. Д. Гогунський // Інформаційні технології в освіті, науці та виробництві. – Вип. 4 (5). — О. : АО Бахва, 2013. — С. 198 — 203.
26. Hirsch, J. E. An index to quantify an individual's scientific research output [Текст] // arXiv: physics/0508025. – v5. – 29 Sep. 2005. – 5 p.
27. Бурков, В. Н. Параметры цитируемости научных публикаций в наукометрических базах данных [Текст] / В. Н. Бурков, А. А. Белощицкий, В. Д. Гогунский // Управління розвитком складних систем. — 2013. — № 15. — С. 134 — 139.
28. Бушуев, С.Д. Напрями дисертаційних наукових досліджень зі спеціальності «Управління проектами та програмами» [Текст] / С.Д. Бушуев, В.Д. Гогунський, К.В. Кошкін // Управління розвитком складних систем. — 2012. – № 12. – С. 5 – 7.
29. Harzing, Anne-Wil. The Publish or Perish Book. – Tarma Software Research Pty Ltd, Мельбурн, Австралія. – 2010. – 266 с. – ISBN 978-0-9808485 – 2–6.
30. Белощицкий, А. А. Управление проблемами в методологии проектно-векторного управления образовательными средами [Текст] / А. А. Белощицкий // Управління розвитком складних систем. – 2012. – № 9. – С. 104 – 107.

31. Чернов, С.К. Концептуальные основы развития наукоемких предприятий в конкурентной среде [Текст] / С.К. Чернов, К. В. Кошкин // Вост.-Европейский журнал передовых технологий. – № 1/2 (43). – Харьков : Технолог. центр, 2010. – С. 20 – 22.
32. Оборський, Г.О. Стандартизація і сертифікація процесів управління якістю освіти у вищому навчальному закладі [Текст] / Г.О. Оборський, В.Д. Гогунський, О.С. Савельєва // Тр. Одес. политехн. ун-та. – Вып. 1(35). – 2011. – С. 251 – 255.
33. Колесникова, Е. В. Управление знаниями в IT-проектах [Текст] / Е. В. Колесникова, А. А. Негри // Вост.-Европ. журнал передовых технологий. – 2013. – № 1/10 (61). – С. 213 – 215.
34. Гогунский, В. Д. Обоснование закона о конкурентных свойствах проектов [Текст] / В. Д. Гогунский, С. В. Руденко, П. А. Тесленко // Управління розвитком складних систем. – № 8. – 2012. – С. 14 – 16.
35. Buneman, P. (1997). [Semistructured data](#). Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems, 117 – 121.
36. Yigal, A., Chin, Y. C., Chun-Nan, H., Craig, A. K. (1993). Retrieving and integrating data from multiple information sources. International Journal of Intelligent and Cooperative Information Systems, vol. 2, issue 2.
37. Jane, Yung-Jen H., Wen-tau, Y. (1997). Template-based information mining from HTML documents. Proceedings of the fourteenth national conference on artificial intelligence and ninth conference on Innovative applications of artificial intelligence, 256 – 262.
38. Dan, S., Mauricio, L. (1997). Information extraction for semi-structured documents. In Proceedings of the Workshop on Management of Semistructured Data.
39. Zhao, L., Wee, K. N., Aixin, S. (2005). Web data extraction based on structural similarity. Journal Knowledge and Information Systems archive, vol. 8, issue 4, 438 – 461.

40. Формализация проблемы извлечения знаний из естественно языковых текстов // [А. Палагин, С. Кривый, Н. Петренко, Д. Бибииков]. — Sofia: Information technologies & knowledge, 2012. — 100 с.
41. Robert Baumgartner. The Personal Publication Reader: Illustrating Web Data Extraction, Personalization and Reasoning for the Semantic Web / Robert Baumgartner, Nicola Henze, Marcus Herzog // Lecture Notes in Computer Science Volume 3532, 2005, pp. 515–530
42. Білощицький, А.О. Ефективність методів пошуку збігів у текстах / А.О. Білощицький, О.В. Діхтяренко // Управління розвитком складних систем. – № 14. – 2013. – С. 144 – 147.
43. Высоцкий, В.Ю. Разработка обучающих программ в виртуальной компьютерной среде // В.Ю. Высоцкий, В.Д. Гогунский // Тр. Одес. политехн. ун-та. – Вып. 2 (36). – 2011. – С. 184 – 189.
44. Яковенко, В.Д. Комп'ютерна реалізація системи автоматизованого управління навчальним процесом // В. Д. Яковенко, В. Д. Гогунський, Г. Ф. Сафонова // Моделир. в прикладных научных исследованиях. Матер. XVI семинара. — Одеса : ОНПУ, 2008. – С. 27 – 30.
45. Тертышная, Т. И. Автоматизированная система контроля знаний / Т. И. Тертышная, Е. В. Колесникова, В. Д. Гогунский // Тр. Одес. политехн. ун-та. — Вып. 1(13).—2001. — С. 125 — 128.
46. Яковенко, А.Е. Стратегия принятия решений в условиях адаптивного обучения / А. Е. Яковенко, А. В. Нарожный, В. Д. Гогунский // Восточно-европейский журнал передових технологий. – № 2/2 (14). – 2005. – С.105 – 110.
47. Чугреев, В. Л. Модель структурного представления текстовой информации и метод ее тематического анализа на основе частотно-контекстной классификации / Санкт-Петербургский гос. электротехнический ун-т "ЛЭТИ" им. В.И. Ульянова. 2003. – С. 25 29.
48. Scott Deerwester, Susan T. Dumais, George W. Furnas, Thomas K.

Landauer, Richard Harshman (1990). Indexing by Latent Semantic Analysis. JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE. 41(6):391–407.

49. Řehůřek, R. (2011). Subspace tracking for latent semantic analysis. Advances in Information Retrieval. 289–300

50. Roger B. Bradford (2008). An empirical study of required dimensionality for large-scale latent semantic indexing applications. In proceeding of: Proceedings of the 17th ACM Conference on Information and Knowledge Management, CIKM 2008, Napa Valley, California, USA.

51. Porter M.F. (1980). An algorithm for suffix stripping. Program, 14 no. 3, pp 130–137.

52. Рач, В. А. Побудова термінологічної системи організації наукового знання [Текст] / В. Рач, О. Россошанська, О. Медведєва // Науковий світ. – 2011. № 4. – С. 13 – 16.

53. Гогунський, В. Д. Марковські моделі комунікаційних процесів в міжнародних проектах / О. В. Власенко, В. В. Лебідь, В. Д. Гогунський // Управління розвитком складних систем. № 12. – 2012. – С. 35 – 39.

54. Плетнев, А.Н. Организация вычислительной сети студгородка «Политехник» с использованием оптического волокна / А.Н. Плетнев, А.Н. Миколук, В.Д. Гогунский // Труды Одес. политехн. ун-та. – № 2(28). – Одеса: ОНПУ, 2007. – С. 138 – 140.

55. Флегантов, Л. Для чего нам нужны международные наукометрические базы данных [Электронный ресурс]. – http://web-in-learning.blogspot.com/2012/11/blog-post_24.html

56. Scopus (Elsevier): [Электронный ресурс]. – <http://www.scopus.com/search/form/authorFreeLookup.url>

57. Web of Science [Электронный ресурс]. – <http://www.webofknowledge.com/>

58. Science Index – Научная электронная библиотека. [Электронный

ресурс]. – <http://www.elibrary.ru>

59. Bielefeld Academic Search Engine. [Электронный ресурс]. – <http://www.base-search.net>

60. Index Copernicus. Indeksacja czasopisma. [Электронный ресурс]. – http://www.journals.indexcopernicus.com/search_article.php

61. Springer Science + Business Media. [Электронный ресурс]. – <http://www.springer.com>

62. Scrapy – a fast high-level screen scraping and web crawling framework. [Электронный ресурс]. – <http://scrapy.org>

63. MongoDB – an open-source document database. [Электронный ресурс]. – <http://ru.wikipedia.org/wiki/MongoDB>

64. Mendeley – Free reference manager and PDF organizer [Электронный ресурс]. – <http://www.mendeley.com>

65. Pedersen, T. Duluth. Word Sense Induction Applied to Web Page Clustering. Proceedings of the 7th International Workshop on Semantic Evaluation (SemEval 2013), in conjunction with the Second Joint Conference on Lexical and Computational Semantics (*SEM-2013). – Atlanta, Georgia, 2013. – P. 202 – 206.

66. Jurgens, D. The S-Space Package: An Open Source Package for Word Space Models / D. Jurgens, K. Stevens // Proceeding ACLDemos '10 Proceedings of the ACL System Demonstrations. – 2010. – P. 30 – 35.

67. Řehůřek, R. (2010). Software Framework for Topic Modelling with Large Corpora / R. Řehůřek, P. Sojka // Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks. – Valletta, Malta, ELRA, 2010. – P. 45 – 50.

68. Agile Planning. [Электронный ресурс]. – http://neilperkin.typepad.com/only_dead_fish/agile-planning.html

69. Lafley A. G., The Game-Changer: How You Can Drive Revenue and Profit Growth with Innovation/ Alan George Lafley Ram Charan. – CrownBusiness, 2008. – 352 p.

70. Cohn M., Agile Estimating and Planning / Mike Cohn. – PrenticeHall,

2005. – 368 p.

71. Agile Manifesto. [Электронный ресурс]. – <http://agilemanifesto.org/>
72. Johnson, S. Where Good Ideas Come From: The Natural History of Innovation / Steven Johnson. – Riverhead Trade, 2011. – 336 p.
73. Воронцов К. В. Вероятностное тематическое моделирование [Электронный ресурс] // www.machinelearning.ru : web. — 2013.
74. Agile Planning. SlideShare. [Электронный ресурс]. – <http://www.slideshare.net/neilperkin/agile-planning-8595549>
75. Risvik K. M., Michelsen R. Search engines and web dynamics (англ.) // Computer Networks : journal. — 2002. — Т. 39. — № 3. — С. 289-302.
76. Segev El. Google and the Digital Divide: The Biases of Online Knowledge. — Oxford: Chandos Publishing. — 2010. — 171 с. — ISBN 978-1-84334-565-7.
77. Tarakeswar M. K., Kavitha M. D. Search Engines: A Study (англ.) // Journal of Computer Applications (JCA) : journal. — 2011. — Т. 4. — № 1. — С. 29-33. — ISSN 0974-1925.
78. Мещеряков В. И., Зайков В. П., Гнатовская А. А. Информационный подход к повышению живучести систем с теплонагруженными приемниками. Восточно-Европейский журнал передовых технологий. Прикладные информационные технологии. – Харьков, 2011. Вып. №3/9 (51). – С. 4 – 7.
79. Зайков, В. П. Влияние термоэлектрической эффективности исходных материалов на показатели надежности термоэлектрических охлаждающих устройств. Часть 1: Однокаскадные ТЭУ [Текст]/ В. П. Зайков, В. И. Мещеряков, А. А. Гнатовская, Ю. И. Журавлев // Технология и конструирование в электронной аппаратуре. – 2015. – № 1. – С. 44 – 48.
80. Палій С. В. Механізми реалізації та впровадження каліграфічного тренажера для дистанційного навчання іноземців / С. В. Палій, М. І. Цюцюра // Управління розвитком складних систем. – 2012. – Вип. 12. – С. 150 – 153.

81. Палій С. В. Соціальні мережі як засіб комунікації електронного навчання / С. В. Палій // Управління розвитком складних систем. – 2013. Вип. 13. – С. 152 – 156.

82. Палій С. В. Користувачі інформаційно-організаційного середовища підготовки іноземців, як джерело знань / С. В. Палій // Системи управління, навігації та зв'язку. – Полтава: ПолтНТУ, 2013. – Вип. 2 (26). – С. 92 – 95.

83. Цюцюра С. В. Застосування нейронних мереж для розпізнавання «ідеального співрозмовника» серед користувачів соціальних мереж / С. В. Цюцюра, І. А. Терейковский, С. В. Палій // Системи управління, навігації та зв'язку. – Полтава: ПолтНТУ, 2013. – Вип. 4 (28). – С. 123 – 126.

ПРИЛОЖЕНИЯ

ПРИЛОЖЕНИЕ А

РЕЗУЛЬТАТЫ ПОИСКА ПУБЛИКАЦИЙ В РАЗРАБОТАННОЙ СИСТЕМЕ
SMD И АНАЛОГЕ – PUBLISH OR PERISH (БАЗА GOOGLE SCHOLAR)

Поиск по атрибуту ФИО автора: Коляда А. С.

Ключевые слова: баз, сист, веб, латентн, научн, scientometr, семантическ, информацион, информац, inform, публикац.

На рис. А.1 показана веб-страница результатов поиска публикаций в информационно-поисковой системе SMD.

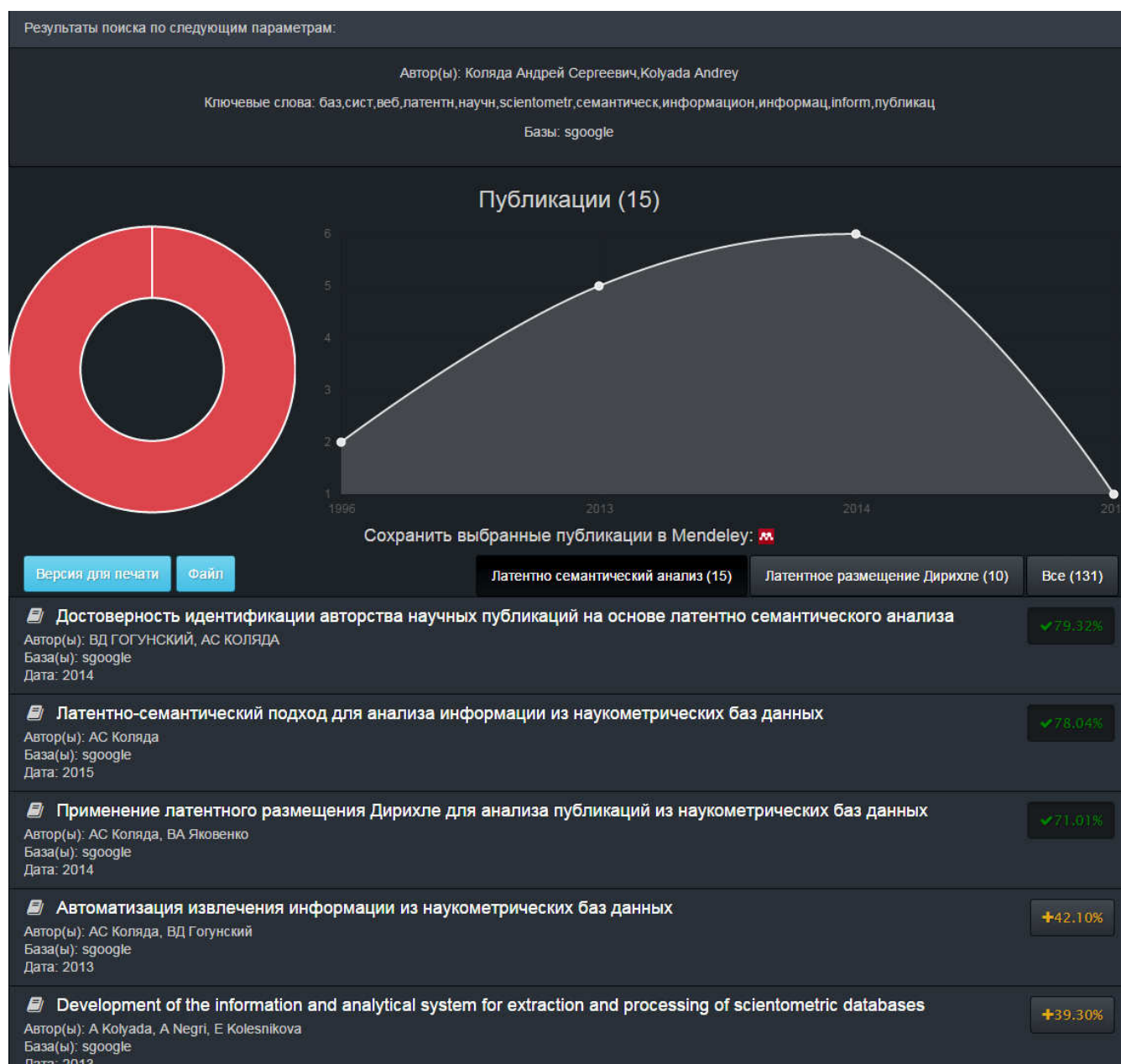


Рисунок А.1 – Результаты поиска публикаций в системе SMD

Найденные публикации с применением ранжирования по LSA

1. Достоверность идентификации авторства научных публикаций на основе латентно семантического анализа (<http://cyberleninka.ru/article/n/dostovernost-identifikatsii-avtorstva-nauchnyh-publikatsiy-na-osnove-latentno-semanticheskogo-analiza>)
ВД ГОГУНСКИЙ, АС КОЛЯДА
2014
Восточно-Европейский журнал
2. Латентно-семантический подход для анализа информации из наукометрических баз данных (<http://journals.uran.ua/index.php/2219-5300/article/view/38753>)
АС Коляда
2015
Управління розвитком складних систем
3. Применение латентного размещения Дирихле для анализа публикаций из наукометрических баз данных (http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Popu_2014_1_32.pdf)
АС Коляда, ВА Яковенко
2014
Праці Одеського
4. Автоматизация извлечения информации из наукометрических баз данных (http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Urss_2013_16_20.pdf)
АС Коляда, ВД Гогунский
2013
Управління розвитком складних
5. Development of the information and analytical system for extraction and processing of scientometric databases
(https://scholar.google.com/scholar?start=120&q=Kolyada&hl=ru&as_sdt=0,5)
А Kolyada, А Negri, Е Kolesnikova
2013
Project management: state and prospects.
6. Разработка проекта информационно-аналитической системы извлечения и обработки информации из наукометрических баз данных
(<http://conference.nuos.edu.ua/catalog/files/lectures/22937.pdf>)
АС Коляда
2013
стан та перспективи. Матеріали ІХ Міжнар.
7. Automating the extraction of information from scientometric databases
(https://scholar.google.com/scholar?start=130&q=Kolyada&hl=ru&as_sdt=0,5)
А Kolyada, V Gogunsky
2013
Management of complex systems
8. Automation scientometric extract information from databases Management of development of difficult systems
(https://scholar.google.com/scholar?start=200&q=Kolyada&hl=ru&as_sdt=0,5)
AS Kolyada, VD Gogunsky
2013
Kyiv;Ukraine: KNUCA: 16;96

9. **Наукометрические данные научного издания «Управление развитием сложных систем»** (<http://urss.knuba.edu.ua/annotation/673>)
ВД Гогунский, АС Коляда
2014
Управління розвитком
10. **Извлечение информации из слабоструктурированных веб-страниц** (<http://cyberleninka.ru/article/n/izvlechenie-informatsii-iz-slabostrukturirovannyh-veb-stranits>)
АС Коляда, ВД Гогунский
2014
Восточно-Европейский журнал
11. **Extracting information from the semistructured web pages** (<http://journals.uran.ua/eejet/article/view/19496>)
АС Коляда, ВД Гогунский
2014
Eastern-European Journal of
12. Elements of the theory and applications of modular technique of parallel information processing. Modern Problems of Optics, Radiation Materials ...
(https://scholar.google.com/scholar?start=130&q=Kolyada&hl=ru&as_sdt=0,5)
АА Kolyada, VV Revinsky, MY Selyaninov
1996
Proc. Sci. Research Inst. Appl. Phys.
13. Рецензенты: к. б. н., старший научный сотрудник Горнотаежной станции ДВО РАН ТА Москалюк
(https://scholar.google.com/scholar?start=320&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)
ВД Фролов, АС Коляда
14. **Application of Latent Dirichlet allocation for the analysis of scientometric publications database** (https://scholar.google.com/scholar?start=140&q=Kolyada&hl=ru&as_sdt=0,5)
VD Gogunsky, VO Iakovenko, AS Kolyada
2014
Proc. of Odes. Polytechnic. Univ. Odessa
15. High-speed methods and systems of digital information processing
(https://scholar.google.com/scholar?start=150&q=Kolyada&hl=ru&as_sdt=0,5)
AF Chernyavski, VV Danilevich, AA Kolyada
1996
Minsk: Belgosuniversitet

Найденные публикации с применением ранжирования по LDA

1. **Достоверность идентификации авторства научных публикаций на основе латентно семантического анализа** (<http://cyberleninka.ru/article/n/dostovernost-identifikatsii-avtorstva-nauchnyh-publikatsiy-na-osnove-latentno-semanticheskogo-analiza>)
ВД ГОГУНСКИЙ, АС КОЛЯДА
2014
Восточно-Европейский журнал
2. **Латентно-семантический подход для анализа информации из наукометрических баз данных** (<http://journals.uran.ua/index.php/2219-5300/article/view/38753>)
АС Коляда
2015
Управління розвитком складних систем

3. Применение латентного размещения Дирихле для анализа публикаций из наукометрических баз данных (http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Popu_2014_1_32.pdf)
АС Коляда, ВА Яковенко
2014
Праці Одеського
4. Автоматизация извлечения информации из наукометрических баз данных (http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Urss_2013_16_20.pdf)
АС Коляда, ВД Гогунский
2013
Управління розвитком складних
5. Development of the information and analytical system for extraction and processing of scientometric databases (https://scholar.google.com/scholar?start=120&q=Kolyada&hl=ru&as_sdt=0,5)
А Kolyada, А Negri, Е Kolesnikova
2013
Project management: state and prospects.
6. Разработка проекта информационно-аналитической системы извлечения и обработки информации из наукометрических баз данных (<http://conference.nuos.edu.ua/catalog/files/lectures/22937.pdf>)
АС Коляда
2013
: стан та перспективи. Матеріали ІХ Міжнар.
7. Automating the extraction of information from scientometric databases (https://scholar.google.com/scholar?start=130&q=Kolyada&hl=ru&as_sdt=0,5)
А Kolyada, V Gogunsky
2013
Management of complex systems
8. Automation scientometric extract information from databases Management of development of difficult systems (https://scholar.google.com/scholar?start=200&q=Kolyada&hl=ru&as_sdt=0,5)
AS Kolyada, VD Gogunsky
2013
Kyiv;Ukraine: KNUCA: 16;96
9. Наукометрические данные научного издания «Управление развитием сложных систем» (<http://urss.knuba.edu.ua/annotation/673>)
ВД Гогунский, АС Коляда
2014
Управління розвитком
10. Извлечение информации из слабоструктурированных веб-страниц (<http://cyberleninka.ru/article/n/izvlechenie-informatsii-iz-slabostrukturirovannyh-veb-stranits>)
АС Коляда, ВД Гогунский
2014
Восточно-Европейский журнал

Все найденные публикации системой SMD

1. Достоверность идентификации авторства научных публикаций на основе латентно семантического анализа (<http://cyberleninka.ru/article/n/dostovernost-identifikatsii-avtorstva-nauchnyh-publikatsiy-na-osnove-latentno-semanticheskogo-analiza>)
ВД ГОГУНСКИЙ, АС КОЛЯДА
2014
Восточно-Европейский журнал
2. Латентно-семантический подход для анализа информации из наукометрических баз данных (<http://journals.uran.ua/index.php/2219-5300/article/view/38753>)
АС Коляда
2015
Управління розвитком складних систем
3. Применение латентного размещения Дирихле для анализа публикаций из наукометрических баз данных (http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DO WNLOAD=1&Image_file_name=PDF/Popu_2014_1_32.pdf)
АС Коляда, ВА Яковенко
2014
Праці Одеського
4. Автоматизация извлечения информации из наукометрических баз данных (http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DO WNLOAD=1&Image_file_name=PDF/Urss_2013_16_20.pdf)
АС Коляда, ВД Гогунский
2013
Управління розвитком складних
5. Development of the information and analytical system for extraction and processing of scientometric databases
(https://scholar.google.com/scholar?start=120&q=Kolyada&hl=ru&as_sdt=0,5)
А Kolyada, А Negri, Е Kolesnikova
2013
Project management: state and prospects.
6. Разработка проекта информационно-аналитической системы извлечения и обработки информации из наукометрических баз данных (<http://conference.nuos.edu.ua/catalog/files/lectures/22937.pdf>)
АС Коляда
2013
: стан та перспективи. Матеріали ІХ Міжнар.
7. Automating the extraction of information from scientometric databases
(https://scholar.google.com/scholar?start=130&q=Kolyada&hl=ru&as_sdt=0,5)
А Kolyada, V Gogunsky
2013
Management of complex systems
8. Automation scientometric extract information from databases Management of development of difficult systems (https://scholar.google.com/scholar?start=200&q=Kolyada&hl=ru&as_sdt=0,5)
AS Kolyada, VD Gogunsky
2013
Kyiv;Ukraine: KNUCA: 16;96
9. Наукометрические данные научного издания «Управление развитием сложных систем» (<http://urss.knuba.edu.ua/annotation/673>)
ВД Гогунский, АС Коляда
2014
Управління розвитком

10. Извлечение информации из слабоструктурированных веб-страниц (<http://cyberleninka.ru/article/n/izvlechenie-informatsii-iz-slabostrukturirovannyh-veb-stranits>)
АС Коляда, ВД Гогунский
2014
Восточно-Европейский журнал
11. Extracting information from the semistructured web pages (<http://journals.uran.ua/eejet/article/view/19496>)
АС Коляда, ВД Гогунский
2014
Eastern-European Journal of
12. Elements of the theory and applications of modular technique of parallel information processing. Modern Problems of Optics, Radiation Materials ... (https://scholar.google.com/scholar?start=130&q=Kolyada&hl=ru&as_sdt=0,5)
АА Кольяда, ВВ Ревинский, МУ Сельянинов
1996
Proc. Sci. Research Inst. Appl. Phys.
13. Рецензенты: к. б. н., старший научный сотрудник Горнотаежной станции ДВО РАН ТА Москалюк (https://scholar.google.com/scholar?start=320&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)
ВД Фролов, АС Коляда
14. Application of Latent Dirichlet allocation for the analysis of scientometric publications database (https://scholar.google.com/scholar?start=140&q=Kolyada&hl=ru&as_sdt=0,5)
VD Gogunsky, VO Iakovenko, AS Kolyada
2014
Proc. of Odes. Polytechnic. Univ. Odessa
15. High-speed methods and systems of digital information processing (https://scholar.google.com/scholar?start=150&q=Kolyada&hl=ru&as_sdt=0,5)
AF Chernyavski, VV Danilevich, AA Kolyada
1996
Minsk: Belgosuniversitet
16. Modular structures of conveyer method of digital information processing (https://scholar.google.com/scholar?start=210&q=Kolyada&hl=ru&as_sdt=0,5)
АА Кольяда, ИТ Пак
1992
Minsk: University
17. "Scientometric data scientific publication" Management of development of difficult systems (https://scholar.google.com/scholar?start=140&q=Kolyada&hl=ru&as_sdt=0,5)
VD Gogunsky, AS Kolyada, VO Iakovenko
2014
of development of difficult systems. Kyiv
18. Использование морфологических признаков в идентификации древесных растений Дальнего Востока в зимний период (https://scholar.google.com/scholar?start=170&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)
АС Коляда, ЕГ Вернигора
Биологические исследования на Горнотаежной
19. HARMFUL CHEMICAL-INDUSTRY INCIDENTS EFFECTS PROGNOSIS SYSTEM OF THE TERRITORIAL CENTER FOR EMERGENCY MEDICINE: SUPPLY OF ... (http://link.springer.com/chapter/10.1007/1-4020-4170-5_19)
АА Кольяда
2006
Treatment of Intoxications and Decontamination of
20. Etching Solution for Developing Aluminum Alloy Welds (https://scholar.google.com/scholar?start=350&q=Kolyada&hl=ru&as_sdt=0,5)
А Кольяда

Patent Number SU 1260716 A(01260716/SU-A)

21. Protection of Power Transmission Lines From Corrosion and Polluted Atmospheres (Zashchita Liniy Elektropredachi ot Korrozii i Zagryaznenosti ...
(https://scholar.google.com/scholar?start=340&q=Kolyada&hl=ru&as_sdt=0,5)

PI Anastasiev, AV Kolyada, EG Proektor
1983

"Energoatomizdat";167

22. Биологические исследования на Горнотаежной станции в период 1932-1940 гг
(https://scholar.google.com/scholar?start=150&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)

АС Коляда
1995

на Горнотаежной станции. Уссурийск: ДВО РАН

23. Radioisotope instrument for measuring the position of interface of two liquid media with similar density (https://inis.iaea.org/search/search.aspx?orig_q=RN:11496985)

VN Afanas'ev, AN Kolyada, II Krejndlin, YI Pakhunkov
1977

24. Transcriptional Regulation of the Human Na⁺/H⁺ Exchanger (NHE-1) Gene
(https://scholar.google.com/scholar?start=380&q=Kolyada&hl=ru&as_sdt=0,5)

AY Kolyada, NE Madias
1997

Blood Purification

25. A novel dimeric inhibitor targeting Beta2GPI in Beta2GPI/antibody complexes implicated in antiphospholipid syndrome (<http://dx.plos.org/10.1371/journal.pone.0015345>)

A Kolyada, CJ Lee, A De Biasio, N Beglova
2010

PLoS One

26. Signal Processing in Noise: A New Methodology
(https://scholar.google.com/scholar?start=50&q=Kolyada&hl=ru&as_sdt=0,5)

VP Tuzlukov, YS Shinakov, AA Kolyada
1998

27. Управління проектами: стан та перспективи
(<http://conference.nuos.edu.ua/catalog/files/lectures/1988.pdf>)

АС Коляда
2013

наук.-практ. конф./АС Коляда; АА Негри; ЕВ

28. Method for protecting renal tubular epithelial cells from radiocontrast nephropathy (RCN)
(<https://www.google.com/patents/US20100003190>)

AY Kolyada, BL Jaber, O Liangos
2007

US Patent App. 12/

29. Этимология названий весенних растений Приморского края (Учебное пособие)
(https://scholar.google.com/scholar?start=150&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)

ВД Фролов, АС Коляда
1996

Уссурийск: Изд-во Уссурийского государственного

30. Regulation of CYP1A induction in hepatoma 27 depending on the site of transplantation
(<http://informahealthcare.com/doi/pdf/10.3109/00498259309059407>)

V Kobliakov, L Kulikova, A Kolyada
1993

31. Seasonal temperature fluctuations in the high northern latitudes during the Cretaceous Period: isotopic evidence from Albian and Coniacian shallow-water ...
(<http://www.sciencedirect.com/science/article/pii/S0195667104001120>)

AM Popov, VV Golozubov, AA Kolyada
2005

Cretaceous

32. Определитель травянистых весеннецветущих растений юга Приморского края (пособие к весенней полевой практике для студентов ...
(https://scholar.google.com/scholar?start=150&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)

ВД Фролов, АС Коляда
1998

Уссурийск: УГПИ

33. Table-type modular arithmetic processors for digital image processing
(<http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=984115>)

А Kolyada, E Otlivanchik
1995

Digital Image

34. Pericyte Rho GTPase mediates both pericyte contractile phenotype and capillary endothelial growth state (<http://www.sciencedirect.com/science/article/pii/S0002944010620019>)

ME Kutcher, AY Kolyada, HK Surks
2007

The American journal of

35. Phenylethanolamine N-methyltransferase gene polymorphisms and adverse outcomes in acute kidney injury (<http://www.karger.com/Article/Fulltext/276577>)

DT O'Connor, MC Perianayagam, AY Kolyada
2010

Nephron Clinical

36. Transcriptional regulation of the human iNOS gene in vascular-smooth-muscle cells and macrophages: evidence for tissue specificity
(<http://www.sciencedirect.com/science/article/pii/S0006291X96904490>)

AY Kolyada, N Savikovsky, NE Madias
1996

Biochemical and biophysical

37. CLAD FERRITIC STEELS

(https://scholar.google.com/scholar?start=400&q=Kolyada&hl=ru&as_sdt=0,5)

LV MEANDROV, VM CHUB, AA KOLYADA
1968

MASHINOSTR;1968,--5--

38. NADPH oxidase p22phox and catalase gene variants are associated with biomarkers of oxidative stress and adverse outcomes in acute renal failure (<http://jasn.asnjournals.org/content/18/1/255.short>)

MC Perianayagam, O Liangos, AY Kolyada
2007

Journal of the

39. Comparative analysis of urinary biomarkers for early detection of acute kidney injury following cardiopulmonary bypass (<http://www.tandfonline.com/doi/abs/10.1080/13547500903067744>)

H Tighiouart, MC Perianayagam, A Kolyada
2009

Biomarkers

40. Cold Rolling of Thick-Walled Rims From Steel 12 KhM
(https://scholar.google.com/scholar?start=480&q=Kolyada&hl=ru&as_sdt=0,5)

VZ Lev, A Kolyada
1978

Khim. Neft. Mashinost.

41. Generation of integral characteristics of symmetric-range residue codes
(<http://www.springerlink.com/index/T81223P174318830.pdf>)

AA Kolyada, MY Selyaninov
1986

Cybernetics and Systems Analysis

42. BUTT WELDING TUBES MADE OF BILAYER METAL KH 14 N 18 V 2 BR PLUS M 3 S COPPER (https://scholar.google.com/scholar?start=350&q=Kolyada&hl=ru&as_sdt=0,5)

IN MEDRISH, VG BENDRIK, AA KOLYADA

1966

KHIM NEFTYANOE

43. STATE OF MONO-OXYGENASE ENZYME-SYSTEM IN REPEATED ADMINISTRATION OF DIETHYL NITROSAMINE (https://scholar.google.com/scholar?start=430&q=Kolyada&hl=ru&as_sdt=0,5)
AY KOLYADA, VA KOBLYAKOV

1980

VOPROSY

44. Localisation of cytochrome P-450PB in the rat liver and primary liver tumors of varying histological structure after the administrations of the inductor ...
(https://scholar.google.com/scholar?start=170&q=Kolyada&hl=ru&as_sdt=0,5)

A Kolyada, V Gelstein, V Kobliakov

1982

Bulletin Experimental Biology and Medicine;10

45. Лекарственные растения Приморья: свойства и применение
(https://scholar.google.com/scholar?start=80&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)

АС Коляда, ВД Фролов

1992

Владивосток: Дальпресс

46. EFFECTS OF HEATING ON THE BRITTLINESS OF KH 25 T STEEL
(https://scholar.google.com/scholar?start=380&q=Kolyada&hl=ru&as_sdt=0,5)

AS ZUBCHENKO, AA KOLYADA, AP ALIKIN

1969

AUTOMAT WELD

47. Monoclonal antibody study of cytochrome P-450 isoforms in the liver of rats treated with phenobarbital, 3-methylcholanthrene, and aroclor 1254
(<http://www.springerlink.com/index/k611143kl1v28076.pdf>)

AY Kolyada

1986

Bulletin of Experimental Biology and Medicine

48. Corrosion and Mechanical Properties of High-Cr Ferritic Vacuum-Cast Steels
(https://scholar.google.com/scholar?start=170&q=Kolyada&hl=ru&as_sdt=0,5)

TB Tokareva, A Kolyada

1977

Zashch. Met.

49. Mechanical and Corrosion Properties of Vacuum-Smelted High-Cr Ferritic Steels
(https://scholar.google.com/scholar?start=340&q=Kolyada&hl=ru&as_sdt=0,5)

TB Tokareva, A Kolyada

1977

Prot. Met.

50. Role of C/EBP proteins in hepatic and vascular smooth muscle transcription of human NHE1 gene
(<http://ajpcell.physiology.org/content/269/6/C1408.short>)

AY Kolyada, CA Johns

1995

American Journal of

51. STATE OF ENZYMATIC SYSTEM OF MICROSOMAL OXYGENASES IN REPEATED ADMINISTRATION OF 2-ACETYL AHINOFLUORENE
(https://scholar.google.com/scholar?start=450&q=Kolyada&hl=ru&as_sdt=0,5)

VA KOBLYAKOV, AY KOLYADA

1978

VOPROSY

52. Radiation source unit for a stationary radiometric level gage
(https://inis.iaea.org/search/search.aspx?orig_q=RN:6218405)

II Krejndlin, AN Kolyada, YI Pakhunkov

1972

Radiation

53. Никольск-Уссурийский: штрихи к портрету
(https://scholar.google.com/scholar?start=10&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)
АС Коляда, АМ Кузнецов
1997
Уссурийск: УГПИ
54. Regulation of endothelial nitric oxide synthase (eNOS) gene by cyclosporine (CsA)
(https://scholar.google.com/scholar?start=220&q=Kolyada&hl=ru&as_sdt=0,5)
АУ Kolyada, NE Madias
1998
J Am Soc Nephrol
55. [Productivity of field crop rotation crops when applying mineral fertilizers periodically and annually [potassium fertilizers, nitrogen fertilizers, ... (<http://agris.fao.org/agris-search/search.do?recordID=SU8000243>)]
VV Turchin, GV Ryabushko, AN Kolyada
1979
Agrokhimiya (USSR
- 56. Authenticity of authorship of scientific publications using latent semantic analysis (<http://journals.uran.ua/eejet/article/view/23942>)**
АС Коляда, ВД Гогунский
2014
Eastern-European Journal of
57. Meridional temperature gradient for Coniacian time, Late Cretaceous (Oxygen and carbon isotopic data on Koryak Upland and Hokkaido)
(https://scholar.google.com/scholar?start=210&q=Kolyada&hl=ru&as_sdt=0,5)
АМ Попов, VV Golozubov, АА Kolyada
2002
The IV International
58. DETERMINATION OF THE INTENSITY OF A STOCHASTIC FLOW GENETICALLY CONNECTED WITH A PRIMARY FLOW
(https://scholar.google.com/scholar?start=410&q=Kolyada&hl=ru&as_sdt=0,5)
AS GAKHOVICH, АА KOLYADA
1979
59. Araliaceae: ginseng and others
(https://scholar.google.com/scholar?start=60&q=Kolyada&hl=ru&as_sdt=0,5)
YN Zhuravlev, AS Kolyada
1996
Dalnauka; Vladivostok
60. VIBRATIONS DECREASE IN MACHINES WITH CONTINUOUS IMPACT ACTION BY INTRODUCTION OF NON-LINEAR RESILIENT ELEMENTS (<http://cyberleninka.ru/article/n/vibrations-decrease-in-machines-with-continuous-impact-action-by-introduction-of-non-linear-resilient-elements>)
61. А KOLYADA, T SOKOL, D PROKOPENKO
2012
Нові матеріали і
62. Transcriptional regulation of the human iNOS gene by IL-1beta in endothelial cells.
(<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950040/>)
АУ Kolyada, NE Madias
2001
Molecular medicine
63. ROLE OF THE C/EBP FAMILY OF PROTEINS IN THE REGULATION OF HEPATIC AND VASCULAR-SMOOTH-MUSCLE TRANSCRIPTION OF THE ...
(https://scholar.google.com/scholar?start=440&q=Kolyada&hl=ru&as_sdt=0,5)
АУ KOLYADA, CA JOHNS
1994

JOURNAL OF

64. Костянка или пиренарий? О типе плода в семействе Araliaceae Juss
(https://scholar.google.com/scholar?start=80&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)

АС Коляда

2000

биологи ческого разнообразия и особенности его

65. К вопросу охраны дальневосточных представителей семейства Araliaceae
(https://scholar.google.com/scholar?start=190&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)

АС Коляда

Материалы конференции, посвященной

66. Araliaceae: Ginseng and Other Aralia of the Russian Far East
(https://scholar.google.com/scholar?start=190&q=Kolyada&hl=ru&as_sdt=0,5)

N Zhuravlev Yu, AS Kolyada

1996

67. Ага/Еасеае: женьшень и другие
(https://scholar.google.com/scholar?start=170&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)

ЮН Журавлев, АС Коляда

1996

68. Cytochrome P-450 content and its induction in primary rat liver tumors
(<http://www.springerlink.com/index/J3758738066608U2.pdf>)

AY Kolyada, VI Gel'shtein, VA Koblyakov

1988

Bulletin of Experimental Biology

69. 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitors upregulate inducible NO synthase expression and activity in vascular smooth muscle cells
(<http://hyper.ahajournals.org/content/38/5/1024.short>)

AY Kolyada, A Fedtsov, NE Madias

2001

Hypertension

70. Sequence of a human liver cytochrome P-450 cDNA clone.
(<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC332242/>)

AY Kolyada

1990

Nucleic acids research

71. Interleukin-8 and acute kidney injury following cardiopulmonary bypass: a prospective cohort study
(<http://www.karger.com/Article/Fulltext/232595>)

Liangos, A Kolyada, H Tighiouart

2009

Nephron Clinical

72. Nikolsk-Ussuriyskiy: shtrikhi k portretu [Nikolsk-Ussuriisk: Strokes to the Portrait]
(https://scholar.google.com/scholar?start=180&q=Kolyada&hl=ru&as_sdt=0,5)

AS Kolyada

1997

Ussuriysk; Izd-vo Ussur. gos. ped. in-ta

73. Nikolsk-Ussuriyskiy: sketches to the portrait
(https://scholar.google.com/scholar?start=180&q=Kolyada&hl=ru&as_sdt=0,5)

AS Kolyada, AM Kuznetsov

1997

74. Лавинный диод от ООО Элемент-Преобразователь (http://power-e.ru/2008_1_10.php)

Р Рыбак, А Коляда

power-e.ru

75. Cluster finding in silicon detectors (<http://elib.bsu.by/handle/123456789/25044>)

AM Kolyada

2012

76. Virus diseases of woody plants. (<http://www.cabdirect.org/abstracts/19950602397.html>)

VD Chernyshev, AS Kolyada, MV Mikhailin

1989

Lesnoe

77. Otsenka bykov po molochnoj produktivnosti docherej. (<http://agris.fao.org/agris-search/search.do?recordID=SU19880015731>)

AF Kolyada

1986

78. Oxygen and carbon isotope composition of Cretaceous organogenic carbonates of Koryak Upland.

Paper 2. Talovka River basin (Koryak Upland)

(https://scholar.google.com/scholar?start=90&q=Kolyada&hl=ru&as_sdt=0,5)

Y Shigeta, H Maeda, TB Afanasyeva, AA Kolyada

2002

Tikhookeanskaya Geologiya

79. Морфология однолетних побегов дальневосточных древесных видов семейства Araliaceae Juss

(https://scholar.google.com/scholar?start=270&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)

AC Коляда

2001

Биологические исследования на Горнотаежной

80. Araliaceae: женьшень и другие

(https://scholar.google.com/scholar?start=10&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)

ЮН Журавлев, AC Коляда

1996

Владивосток: Дальнаука

81. Cyclosporine a (CsA) inhibits the transcriptional activity of the human inducible nitric oxide synthase (iNOS) promoter in vascular-smooth-muscle (...

(https://scholar.google.com/scholar?start=420&q=Kolyada&hl=ru&as_sdt=0,5)

AY Kolyada, NE Madias

1996

JOURNAL OF

82. Statisticheskiy analiz sluchaynykh potokov v fizicheskom eksperimente [Statistical analysis of stochastic flows in physical experiment]

(https://scholar.google.com/scholar?start=180&q=Kolyada&hl=ru&as_sdt=0,5)

VV Apanasovich, AA Kolyada, AF Chernyavsky

1988

Minsk: Universitetskoe Publ

83. Failure of titanium-clad 18-8 chromium-nickel steel coils in melamine production

(<http://www.springerlink.com/index/J17L753838018756.pdf>)

AS Zubchenko, AA Kolyada, VL Shaleeva

1969

Chemical and Petroleum

84. Probing Anticoagulant Fondaparinux for Interference with Prothrombotic B2GPI/Anti-B2GPI Antibody Complexes (https://scholar.google.com/scholar?start=390&q=Kolyada&hl=ru&as_sdt=0,5)

A Kolyada, A De Biasio, N Beglova

2011

BLOOD

85. INFLUENCE OF NITROGEN ON COLD BRITTLENESS OF HEAT-RESISTANT STEEL KH 25

YU 5 (https://scholar.google.com/scholar?start=420&q=Kolyada&hl=ru&as_sdt=0,5)

MM TIMOFEEV, EL KAZIMIROVSKAYA, AA KOLYADA

1970

AVTOMAT SVARKA;SEPT.

86. IMMUNOCHEMICAL IDENTIFICATION OF 5 ISOFORMS OF CYTOCHROME-P-450 BY IMMUNOBLOTTING WITH MONOCLONAL-ANTIBODIES
(https://scholar.google.com/scholar?start=220&q=Kolyada&hl=ru&as_sdt=0,5)
AY Kolyada
1987
BIOLOGICHESKIE MEMBRANY
87. Effect of C and N on the Weldability of Ferritic Cr Steels
(https://scholar.google.com/scholar?start=110&q=Kolyada&hl=ru&as_sdt=0,5)
AS Zubchenko, A Kolyada
1972
Avtomat. Svarka
88. EFFECTS OF NITROGEN ON THE COLD SHORTNESS OF THE HEAT RESISTANT KH 25 YU 5 STEEL (https://scholar.google.com/scholar?start=340&q=Kolyada&hl=ru&as_sdt=0,5)
MM TIMOFEEV, EL KAZIMORIVSKAYA, AA KOLYADA
1970
AUTOMAT WELD
89. Statistical Analysis of Random Flows in a Physical Experiment
(https://scholar.google.com/scholar?start=70&q=Kolyada&hl=ru&as_sdt=0,5)
VV Apanasovich, AA Kolyada, AF Chernyavskii
1988
Universitetskoe;Minsk
90. Effect of N on the Properties of Low-C Ferritic Steel With 25 per cent Cr
(https://scholar.google.com/scholar?start=350&q=Kolyada&hl=ru&as_sdt=0,5)
IR Patskevich, AS Zubchenko, A Kolyada
1975
Met. Sci. Heat Treat.
91. Proximal regulatory elements and nuclear activities required for transcription of the human Na⁺/H⁺ exchanger (NHE-1) gene (<http://www.sciencedirect.com/science/article/pii/0167478194901244>)
AY Kolyada, TV Lebedeva, CA Johns
1994
Biochimica et Biophysica
92. Algorithms in generalized RCS arithmetic
(https://scholar.google.com/scholar?start=200&q=Kolyada&hl=ru&as_sdt=0,5)
AA Kolyada
1980
Vestn. Beloruss. Univ.;Ser
93. EFFECTS OF CARBON AND NITROGEN ON WELDABILITY OF FERRITIC CHROMIUM STEELS (https://scholar.google.com/scholar?start=450&q=Kolyada&hl=ru&as_sdt=0,5)
ZUBCHENK. AS, AA KOLYADA
1972
AUTOMATIC
94. A genetic variant of hypoxia-inducible factor-1 α is associated with adverse outcomes in acute kidney injury (<http://www.nature.com/ki/journal/v75/n12/abs/ki200968a.html>)
AY Kolyada, H Tighiouart, MC Perianayagam
2009
Kidney
95. O-TETRAHYDROFURANYL DERIVATIVES OF THYMIDINE AND 2'-DEOXYCYTIDINE
(https://scholar.google.com/scholar?start=150&q=Kolyada&hl=ru&as_sdt=0,5)
GA BELITSKY, AY Kolyada
1980
96. Hypoxia-Inducible Factor-1 α Stabilizing Genetic Variant and Adverse Outcomes in Acute Kidney Injury (<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2872511/>)
AY Kolyada, H Tighiouart, MC Perianayagam
2009
Kidney

97. Effect of chromium and nickel on the corrosion resistance of the metal of welded joints in pyrolysis coils (<http://www.springerlink.com/index/rv01742u71p51768.pdf>)
AA Kolyada, AG Parkhomenko
1986
Chemical and Petroleum Engineering
98. Immunohistochemical localization of cytochrome P-450 in rat liver during phenobarbital induction (<http://www.springerlink.com/index/M5W147787RU68427.pdf>)
AY Kolyada
1981
Bulletin of Experimental Biology and Medicine
99. Effect of nitrogen on the properties of low-carbon ferritic steel with 25% Cr (<http://www.springerlink.com/index/W7T617032K924472.pdf>)
IR Patskevich, AS Zubchenko, AA Kolyada
1975
Metal Science and Heat
100. Особенности структуры покровных тканей стволов древесных растений Приморского края (https://scholar.google.com/scholar?start=180&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)
АС Коляда
Животный и растительный мир Дальнего Востока
101. . КЛАССИФИКАЦИЯ ПЛОДОВ ДРЕВЕСНЫХ РАСТЕНИЙ ПРИМОРСКОГО КРАЯ (<http://cyberleninka.ru/article/n/klassifikatsiya-plodov-drevesnyh-rasteniy-primorskogo-kрая>)
АС Коляда, НА Коляда
2009
Вестник Оренбургского
102. . О ПУТЯХ РАСПРОСТРАНЕНИЯ ДИАСПОР ДРЕВЕСНЫХ РАСТЕНИЙ ПРИМОРСКОГО КРАЯ (<http://elibrary.ru/item.asp?id=11735272>)
АС Коляда, НА Коляда
2008
Вестник Красноярского государственного
103. . О ПЕРИДЕРМАЛЬНЫХ ДРЕВЕСНЫХ РАСТЕНИЯХ ЮГА ПРИМОРЬЯ (<http://botsad.ru/media/oldfiles/journal/number5/13.pdf>)
АС Коляда
botsad.ru
104. . HMG CoA reductase inhibitors induce apoptosis in rat aortic vascular-smooth-muscle (VSM) cells. (https://scholar.google.com/scholar?start=130&q=Kolyada&hl=ru&as_sdt=0,5)
AY Kolyada, NE Madias
1996
JOURNAL OF
105. . Plasma cystatin C and acute kidney injury after cardiopulmonary bypass (<http://cjasn.asnjournals.org/content/5/8/1373.short>)
Liangos, MC Perianayagam, A Kolyada
2010
Clinical Journal of the
106. STRATIGRAPHIC SUBDIVISION OF RHYTHMIC TERRIGENOUS DEPOSITS OF CENTRAL PART OF KORYAK HIGHLANDS (https://scholar.google.com/scholar?start=180&q=Kolyada&hl=ru&as_sdt=0,5)
BV Yermakov, AA Kolyada, LV Vasil'-yenko, VF Mishin
1974
AND LITHOLOGY OF
107. . Древесные растения Уссурийского городского округа (http://uss.dvfu.ru/files/publications/kolyada_as_zhivotnyi_i_rastitel_mir_dv_n24_2015.pdf#page=31)
ЯА Гринченко, АС Коляда
ББК 28.0 (2Р55) Ж 67 Составитель

108. . Что скрыто в именах растений
(https://scholar.google.com/scholar?start=70&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)
АС Коляда, ВД Фролов
1998
Востока.(Пособие для учителей). Уссурийск: Изд-во
109. . Интенсивность транспирации и содержание воды в листьях растений в горах Сихоте-Алиня
(https://scholar.google.com/scholar?start=70&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)
ВД Чернышев, АС Коляда
1982
Бот. журн
110. . Butt welding of pipe made of the double-layer metal Kh14N18V2BR steel+ M3S copper
(<http://www.springerlink.com/index/th849756643x1537.pdf>)
IN Medrish, VG Bendrik, AA Kolyada
1966
Chemical and Petroleum
111. . Activation of transcription of the human iNOS and eNOS genes in cultured VSM and endothelial cells by simvastatin. (https://scholar.google.com/scholar?start=460&q=Kolyada&hl=ru&as_sdt=0,5)
AY Kolyada, NE Madias
1997
JOURNAL OF
112. . Smolin, VV; Medvedev, EA Corrosion mechanical properties of high-chromium content ferritic steels produced by vacuum melting (in Russian) ...
(https://scholar.google.com/scholar?start=180&q=Kolyada&hl=ru&as_sdt=0,5)
TB Tokareva, AA Kolyada
1977
113. . Increased plasma IL-8 level is associated with IL-8-251 AA genotype and with acute kidney injury following cardiopulmonary bypass
(https://scholar.google.com/scholar?start=170&q=Kolyada&hl=ru&as_sdt=0,5)
Liangos, A Kolyada, MC Perianayagam, H Tighiouart
2007
J Am Soc Nephrol A
114. . CORROSION RESISTANCE OF CR AND CR-NI STEELS AND THEIR JOINTS IN A MELT OF SODIUM SULFIDE (https://scholar.google.com/scholar?start=480&q=Kolyada&hl=ru&as_sdt=0,5)
AS ZUBCHENKO, AA KOLYADA, PI DAVIDCHUK
1971
AVTOMAT SVARKA;APR.
115. RESEARCH INTO BRITTLE FAILURE TENDENCY OF FERRITIC STEEL KH25T
(https://scholar.google.com/scholar?start=430&q=Kolyada&hl=ru&as_sdt=0,5)
ZUBCHENK. AS, AA KOLYADA
1971
WELDING
116. Regular and chaotic motions of a body over a periodic surface
(<http://www.springerlink.com/index/l486751623110j48.pdf>)
AF Kolyada
2001
International applied mechanics
117. Radioisotope tracking level gage-regulator
(https://inis.iaea.org/search/search.aspx?orig_q=RN:9355105)
II Krejndlin, AN Kolyada, YI Pakhunkov, KS Furman
1974
118. Reciprocal cross differences in Drosophila melanogaster longevity: an evidence for non-genomic effects in heterosis phenomenon? (<http://link.springer.com/article/10.1007/s10522-013-9419-6>)
AM Vaiserman, OG Zabuga, AK Kolyada, AV Pisaruk

2013

Biogerontology

119. О происхождении и синонимах русских названий птиц Приморья (https://scholar.google.com/scholar?start=480&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)

АС Коляда, ЮН Глущенко

Animals and plants of the Russian Far East

120. The effect of alloying on joint properties in welding grade Kh25T steel (<http://www.springerlink.com/index/HU2KJ36487525T58.pdf>)

AS Zubehenko, AA Kolyada, AP Alikin

1969

Chemical and Petroleum

121. Predisposition to type II diabetes among those residents of Ukraine whose prenatal development coincided with the famine of 1932–1933 (<http://link.springer.com/article/10.1134/S2079057011040163>)

AV Pisaruk, LV Mekhova, AK Kolyada

2011

Advances in

122. Changes in the cytochrome P-450 content during carcinogenesis and in hepatomas in rats (https://scholar.google.com/scholar?start=220&q=Kolyada&hl=ru&as_sdt=0,5)

AY Kolyada

1982

Author's Abstract of Candidate's Dissertation;Moscow

123. Rho GTPase signaling modulates cell shape and contractile phenotype in an isoactin-specific manner (<http://ajpcell.physiology.org/content/285/5/C1116.short>)

AY Kolyada, KN Riley

2003

American Journal of

124. MEMBERS OF THE C/EBP FAMILY OF PROTEINS ARE POTENT TRANSACTIVATORS OF THE HUMAN NA⁺/H⁺ EXCHANGER (NHE-1) GENE (https://scholar.google.com/scholar?start=460&q=Kolyada&hl=ru&as_sdt=0,5)

AY KOLYADA, NE MADIAS

1995

JOURNAL OF

125. Protective effect of erythropoietin against radiocontrast-induced renal tubular epithelial cell injury (<http://www.karger.com/Article/Fulltext/110089>)

AY Kolyada, O Liangos, NE Madias

2008

American journal of

126. . Руководство по определению древесных растений Приморского края в зимний период (пособие для преподавателей биологии растений ... (https://scholar.google.com/scholar?start=80&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)

АС Коляда, ВД Фролов

2003

Усурийск: Изд-во УГПИ

127. . COLD CRACKING OF FERRITIC STEEL WELDS (https://scholar.google.com/scholar?start=410&q=Kolyada&hl=ru&as_sdt=0,5)

ММ ТИМОФЕЕВ, АС ЗУБЧЕНКО, АА КОЛЯДА

1970

АВТОМАТ СВАРКА;ОСТ.

128. . Inhibition of thrombotic properties of persistent autoimmune anti-β2GPI antibodies in the mouse model of antiphospholipid syndrome (<http://www.bloodjournal.org/content/123/7/1090.short>)

A Kolyada, A Porter, N Beglova

2014

Blood

129. . Properties and use of two-layer steel Vst. 3+ 0Kh17T (<http://www.springerlink.com/index/un1024834618526m.pdf>)

AA Bykov, SA Gladyshevskaya, AA Kolyada
1970

Chemical and Petroleum

130. . Araliaceae: zhenshei i drugie (<http://kbd.kew.org/kbd/detailedresult.do?id=338494>)

AS Kolyada
1998

: ginseng and others.) Vladivostok: Dalnauka 279p.-

131. . The disturbances in regulation of microsomal enzyme activities by lipids during chemical hepatocarcinogenesis (https://scholar.google.com/scholar?start=150&q=Kolyada&hl=ru&as_sdt=0,5)

MI Dzalyabova, VA Koblyakov, AY Kolyada
1978

Doklady Akad. Nauk SSSR

132. . Полевой определитель основных древесных растений Приморья в зимний период (https://scholar.google.com/scholar?start=180&q=%D0%9A%D0%BE%D0%BB%D1%8F%D0%B4%D0%B0&hl=ru&as_sdt=0,5)

АС Коляда
2009

Учебное пособие. Уссурийск: Изд-во УГПИ

На рис. Б.1 показан скриншот найденных публикации в системе PUBLISH OR PERISH.

Cites	Per year	Rank	Authors	Title	Year	Publication	Publisher	Type
38	1.46	4	АН УССР, АН Шарковский, ...	Динамика одномерных отображений	1989		ir.nmu.org.ua	CITATION
21	30.50	1	АС Коляда, ВД Гогунский	Автоматизация извлечения информации из неметрических баз данных	2013	Управление развитием обл...	irbis-nbuv.gov.ua	PDF
12	0.63	2	ЮН Журавлев, АС Коляда	Ага-Басова: женьшень и другие	1998	Владивосток: Дальнаука		CITATION
9	0.50	3	АС Коляда, АМ Кузнецов	Николай-Уссурийский: штрихи к портрету	1997	Уссурийск: УГПИ		CITATION
6	3.00	1	АС Коляда	Разработка проекта информационно-аналитической системы выявления и обработки информации из неметрич...	2012	... стан та перспектив...	conf.ifp.gov.lvzoo.edu.ua	PDF
6	6.00	6	ВД Гогунский, АС Коляда	Неметрические данные научного издания: управление развитием сложных систем	2014	Управление развитием обл...	irbis-nbuv.gov.ua	CITATION
5	5.00	7	АС Коляда, ВД Гогунский	Извлечение информации из слабоструктурированных веб-страниц	2014	Восточно-Европейский жу...	cyberleninka.ru	HTML
4	0.33	12	АС Коляда, ВД Фролов	Ринкозоства по определению древесных растений Приморского края в зимний период (пособие для преподавател...	2003	Уссурийск: Изд-во УГПИ		CITATION
4	0.44	20	СА Коляда	Подарками как канал распространения преступности несправедливо...	2006	Человек: преступление и ...	elibrary.ru	CITATION
3	0.38	9	АС Коляда, ВД Фролов	Что одоро в янечох растений	1998	Восток (Пособие для у...		CITATION
3	0.08	1	ВД Черышев, АС Коляда	Интенсивность транспирации и содержание воды в листьях растений в горах Сихотэ-Алиня	1982	Bot. журн.		CITATION
3	0.20	20	АС Коляда	Костянка или перхоть? О типе пло-да в семействах Agavaceae, Juss	2000	Биологического рекон...		CITATION
3	0.12	11	АС Коляда, ВД Фролов	Лекарственные растения Приморья: свойства и применение	1992	Владивосток: Дальпресс		CITATION
3	0.25	13	АВ Сукарев, МС Радчинова...	Методические рекомендации по этнофункциональной популяционной реконструкции и восстановлен...	2003	Журнал прикладной психо...		CITATION
3	3.00	14	АС Коляда, ВА Яковина	Применение латентного развития: Директе для анализа публикаций из неметрических баз данных	2014	Праця Східноказа...	irbis-nbuv.gov.ua	PDF
3	3.00	19	ВД ГОГУНСКИЙ, АС КОЛЯДА	Достоверность идентификации авторов научных публикаций на основе латентно-смантического анализа	2014	Восточно-Европейский жу...	cyberleninka.ru	HTML
3	3.00	23	АС Коляда	Латентно-смантический подход для анализа информации из неметрических баз данных	2015	Управление развитием обл...	ryabiy.digitallib.ru	CITATION
2	0.11	15	ВД Фролов, АС Коляда	Этиология названий древесных растений Приморского края (учебное пособие)	1996	Уссурийск: Изд-во Усс...		CITATION
2	0.33	36	АС Коляда	Биологическое исследование на Герметической станции в период 1932-1941 гг	1995	на Герметической станц...		CITATION
2	0.12	17	ВД Фролов, АС Коляда	Определитель травянистых бореальных растений юга Приморского края (пособие к учебной полевой прес...	1998	Уссурийск: УГПИ		CITATION
2	0.09	19	АС Коляда, ВГ Бернгард	Использование морфологических признаков в идентификации древесных растений Дальнего Востока в зимней пер...		Биологические исследова...		CITATION
2	0.11	20	ЮН Журавлев, АС Коляда	Ага-Басова: женьшень и другие	1996		Владивосток	CITATION
2	3.00	21	АС Коляда	Управление проектами: стан та перспектив	2012	матеріал IX міжнар. наук...	conf.ifp.gov.lvzoo.edu.ua	CITATION
2	0.09	22	АС Коляда	Особенности структуры покровных тканей стеблей древесных растений Приморского края		Животный и растительный...		CITATION
2	0.22	23	АС Коляда	Полевой определитель основных древесных растений Приморья в зимний период	2009	Учебное пособие. Уссурий...		CITATION
2	0.00	24	АС Коляда	К вопросу охраны дальневосточных представителей семейства Agavaceae		Материалы конференци...		CITATION
2	0.28	26	... ЮЛ Попов, ВЕ Машук...	Профилактическое лечение эпилепсии: возможности кардиона (нейролеукофармакологического и биоэлектрического и...	2008	...	geopharm.ru	PDF
2	0.67	27	... НО Кендзюров, АС Мельни...	Таксономический анализ популяций дуба в Приморском крае (на основе митохондриальной ДНК)	2012	Вісник ННТУ України	irbis-nbuv.gov.ua	PDF
2	0.08	31	ПН Коляда, СА Царва, АВ Б...	Механизмы реализации в реалити эволюционно-экологического показателя на островах галапагоз у гибридных пш...	1990	Лабораторные заметные...		CITATION
2	0.00	32	ВН Астафуров, АА Вакин, ...	Способ получения эвтаназического раствора четырехвалентного марганца		Лет. свид. на изобретение		CITATION
1	3.00	29	ВД Гогунский, АС Коляда	Неметрические базы характеристик, показателя	2014	Наук. зап., запечатл...	storage.library.ox.ac.uk	PDF
1	0.07	30	АС Коляда	Морфология однолетних побегов дальневосточных древесных видов семейства Agavaceae, Juss	2001	Биологические исследова...		CITATION
1	0.11	11	МП Чепелев, ВП Чепелева	Видовый состав и составитель Дальнего Востока Дальний Т. Палеонтологический ...	2012	Вестн ...	vestnik.dv.ru	PDF

Рисунок Б.1 – Результаты поиска публикаций в системе PUBLISH OR PERISH

Все найденные публикации системой PUBLISH OR PERISH

1. 38,"АН УССР, АН Шарковский, СФ Коляда, АГ Сивак...", "Динамика одномерных отображений", 1989, "", "ir.nmu.org.ua", "", "http://scholar.google.com/scholar?cites=9588698502754636256&as_sdt=2005&scioldt=0,5&hl=en&num=20", 4, 2015-08-20, "CITATION"

2. 21, "АС Коляда, ВД Гогунский", "Автоматизация извлечения информации из наукометрических баз данных", 2013, "Управління розвитком складних ...", "irbis-nbuv.gov.ua", "", "", "http://scholar.google.com/scholar?cites=11769897368675941833&as_sdt=2005&scioldt=0,5&hl=en&num=20", 1, 2015-08-20, "PDF"
3. 12, "ЮН Журавлев, АС Коляда", "Araliaceae: женьшень и другие", 1996, "Владивосток: Дальнаука", "", "", "http://scholar.google.com/scholar?cites=9868159043999839540&as_sdt=2005&scioldt=0,5&hl=en&num=20", 2, 2015-08-20, "CITATION"
4. 9, "АС Коляда, АМ Кузнецов", "Никольск-Уссурийский: штрихи к портрету", 1997, "Уссурийск: УГПИ", "", "", "http://scholar.google.com/scholar?cites=6413731919639684945&as_sdt=2005&scioldt=0,5&hl=en&num=20", 3, 2015-08-20, "CITATION"
5. 6, "АС Коляда", "Разработка проекта информационно-аналитической системы извлечения и обработки информации из наукометрических баз данных", 2013, "... : стан та перспективи. Матеріали ІХ Міжнар. ...", "conference.nuos.edu.ua", "", "", "http://scholar.google.com/scholar?cites=783914820087359832&as_sdt=2005&scioldt=0,5&hl=en&num=20", 5, 2015-08-20, "PDF"
6. 6, "ВД Гогунский, АС Коляда...", "Наукометрические данные научного издания «Управление развитием сложных систем»", 2014, "Управління розвитком ...", "urss.knuba.edu.ua", "<http://urss.knuba.edu.ua/annotation/673>", "http://scholar.google.com/scholar?cites=12396708594821375054&as_sdt=2005&scioldt=0,5&hl=en&num=20", 6, 2015-08-20, ""
7. 5, "АС Коляда, ВД Гогунский", "Извлечение информации из слабоструктурированных веб-страниц", 2014, "Восточно-Европейский журнал ...", "cyberleninka.ru", "", "", "http://scholar.google.com/scholar?cites=9375044799029107467&as_sdt=2005&scioldt=0,5&hl=en&num=20", 7, 2015-08-20, "HTML"
8. 4, "АС Коляда, ВД Фролов", "Руководство по определению древесных растений Приморского края в зимний период (пособие для преподавателей биологии растений средней ...)", 2003, "Уссурийск: Изд-во УГПИ", "", "", "http://scholar.google.com/scholar?cites=11200904119012245258&as_sdt=2005&scioldt=0,5&hl=en&num=20", 12, 2015-08-20, "CITATION"
9. 4, "СА Коляда", "Подражание как канал воспроизводства преступности несовершеннолетних", 2006, "Человек: преступление и наказание", "elibrary.ru", "", "", "http://scholar.google.com/scholar?cites=13245080522198818503&as_sdt=2005&scioldt=0,5&hl=en&num=20", 28, 2015-08-20, "CITATION"
10. 3, "АС Коляда, ВД Фролов", "Что скрыто в именах растений", 1998, "... Востока. (Пособие для учителей). Уссурийск: Изд-во ...", "", "", "http://scholar.google.com/scholar?cites=15125878362650454452&as_sdt=2005&scioldt=0,5&hl=en&num=20", 8, 2015-08-20, "CITATION"
11. 3, "ВД Чернышев, АС Коляда", "Интенсивность транспирации и содержание воды в листьях растений в горах Сихоте-Алиня", 1982, "Бот. журн", "", "", "http://scholar.google.com/scholar?cites=15001472367634812454&as_sdt=2005&scioldt=0,5&hl=en&num=20", 9, 2015-08-20, "CITATION"
12. 3, "АС Коляда", "Костянка или пиренарий? О типе пло да в семействе Araliaceae Juss", 2000, "... биологи чешского разнообразия и особенности его ...", "", "", "http://scholar.google.com/scholar?cites=13277416854137831138&as_sdt=2005&scioldt=0,5&hl=en&num=20", 10, 2015-08-20, "CITATION"
13. 3, "АС Коляда, ВД Фролов", "Лекарственные растения Приморья: свойства и применение", 1992, "Владивосток: Дальпресс", "", "", "http://scholar.google.com/scholar?cites=10447819484592332304&as_sdt=2005&scioldt=0,5&hl=en&num=20", 11, 2015-08-20, "CITATION"
14. 3, "АВ Сухарев, МС Радионова, АК Коляда", "Методические рекомендации по этнофункциональной психопрофилактике психической дезадаптации у воспитанников детских садов", 2003, "Журнал прикладной психологии", "", "", "http://scholar.google.com/scholar?cites=16944686055356495104&as_sdt=2005&scioldt=0,5&hl=en&num=20", 13, 2015-08-20, "CITATION"
15. 3, "АС Коляда, ВА Яковенко...", "Применение латентного размещения Дирихле для анализа публикаций из наукометрических баз данных", 2014, "Праці Одеського ...", "irbis-

- nbuv.gov.ua", "", "http://scholar.google.com/scholar?cites=15852055887480276124&as_sdt=2005&scioldt=0,5&hl=en&num=20",14,2015-08-20,"PDF"
16. 3,"ВД ГОГУНСКИЙ, АС КОЛЯДА","Достоверность идентификации авторства научных публикаций на основе латентно семантического анализа",2014,"Восточно-Европейский журнал ...", "cyberleninka.ru", "", "http://scholar.google.com/scholar?cites=16043651585386683343&as_sdt=2005&scioldt=0,5&hl=en&num=20",18,2015-08-20,"HTML"
17. 3,"АС Коляда","Латентно-семантический подход для анализа информации из наукометрических баз данных",2015,"Управління розвитком складних систем", "journals.uran.ua", "http://journals.uran.ua/index.php/2219-5300/article/view/38753", "http://scholar.google.com/scholar?cites=3456331559207181275&as_sdt=2005&scioldt=0,5&hl=en&num=20",25,2015-08-20,""
18. 2,"ВД Фролов, АС Коляда","Этимология названий весенних растений Приморского края (Учебное пособие)",1996,"Уссурийск: Изд-во Уссурийского государственного ...", "", "", "http://scholar.google.com/scholar?cites=3164236032505386694&as_sdt=2005&scioldt=0,5&hl=en&num=20",15,2015-08-20,"CITATION"
19. 2,"АС Коляда","Биологические исследования на Горнотаежной станции в период 1932-1940 гг.",1995,"... на Горнотаежной станции. Уссурийск: ДВО РАН", "", "", "http://scholar.google.com/scholar?cites=1830307245809155648&as_sdt=2005&scioldt=0,5&hl=en&num=20",16,2015-08-20,"CITATION"
20. 2,"ВД Фролов, АС Коляда","Определитель травянистых весеннецветущих растений юга Приморского края (пособие к весенней полевой практике для студентов высших ...",1998,"Уссурийск: УГПИ", "", "", "http://scholar.google.com/scholar?cites=8701764983985448951&as_sdt=2005&scioldt=0,5&hl=en&num=20",17,2015-08-20,"CITATION"
21. 2,"АС Коляда, ЕГ Вернигора","Использование морфологических признаков в идентификации древесных растений Дальнего Востока в зимний период",0,"Биологические исследования на Горнотаежной ...", "", "", "http://scholar.google.com/scholar?cites=13258994541725544405&as_sdt=2005&scioldt=0,5&hl=en&num=20",19,2015-08-20,"CITATION"
22. 2,"ЮН Журавлсв, АС Коляда","Ага/Еасеае: женьшень и другие",1996,"", "Владивосток", "", "http://scholar.google.com/scholar?cites=3496546204543529671&as_sdt=2005&scioldt=0,5&hl=en&num=20",20,2015-08-20,"CITATION"
23. 2,"АС Коляда","Управління проектами: стан та перспективи",2013,"матеріали ІХ міжнар. наук.-практ. конф./АС ...", "conference.nuos.edu.ua", "http://conference.nuos.edu.ua/catalog/files/lectures/1988.pdf", "http://scholar.google.com/scholar?cites=8655433214885169649&as_sdt=2005&scioldt=0,5&hl=en&num=20",21,2015-08-20,""
24. 2,"АС Коляда","Особенности структуры покровных тканей стволов древесных растений Приморского края",0,"Животный и растительный мир Дальнего Востока", "", "", "http://scholar.google.com/scholar?cites=11790171927028494801&as_sdt=2005&scioldt=0,5&hl=en&num=20",22,2015-08-20,"CITATION"
25. 2,"АС Коляда","Полевой определитель основных древесных растений Приморья в зимний период",2009,"Учебное пособие. Уссурийск: Изд-во УГПИ", "", "", "http://scholar.google.com/scholar?cites=17571384528204778766&as_sdt=2005&scioldt=0,5&hl=en&num=20",23,2015-08-20,"CITATION"
26. 2,"АС Коляда","К вопросу охраны дальневосточных представителей семейства Araliaceae",0,"Материалы конференции, посвященной", "", "", "http://scholar.google.com/scholar?cites=1504912202299933730&as_sdt=2005&scioldt=0,5&hl=en&num=20",24,2015-08-20,"CITATION"
27. 2,"...", ЮИ Поляков, ВЕ Машукова, АА Коляда,"Профилактическое лечение эпилепсии: возможности кортексина (нейроиммунофизиологическое и биохимическое исследование)",2008,"...", "geropharm.ru", "", "http://scholar.google.com/scholar?cites=9532030755299876337&as_sdt=2005&scioldt=0,5&hl=en&num=20",26,2015-08-20,"PDF"
28. 2,"...", НЗ Кендзьора, АС Мельник, ЛБ Коляда,"Таксономічний склад голонасінних дендропарку Львівської клінічної інфекційної лікарні",2012,"... вісник НЛТУ України", "irbis-

- nbuv.gov.ua", "", "http://scholar.google.com/scholar?cites=3060243955169258419&as_sdt=2005&scioldt=0,5&hl=en&num=20", 27, 2015-08-20, "PDF"
29. 2, "ТИ Коляда, СА Хорева, АВ Баталина", "Межлинейные различия в реакции иммунноэндокринологических показателей на острую гипоксию у инбредных мышей", 1990, "Лабораторные животные для медико-... ", "", "", "http://scholar.google.com/scholar?cites=1666581230033823695&as_sdt=2005&scioldt=0,5&hl=en&num=20", 31, 2015-08-20, "CITATION"
30. 2, "ВИ Астафуров, АА Вашман, НС Коляда", "Способ получения азотнокислого раствора четырехвалентного марганца", 0, "Авт. свид. на изобретение", "", "", "http://scholar.google.com/scholar?cites=8515120940729770049&as_sdt=2005&scioldt=0,5&hl=en&num=20", 32, 2015-08-20, "CITATION"
31. 1, "ВД Гогунський, АС Коляда...", "Наукометричні бази: характеристика, можливості і завдання", 2014, "...
Наук.-метод. семінар.
...", "storage.library.opu.ua", "", "http://scholar.google.com/scholar?cites=1786246745824922745&as_sdt=2005&scioldt=0,5&hl=en&num=20", 29, 2015-08-20, "PDF"
32. 1, "АС Коляда", "Морфология однолетних побегов дальневосточных древесных видов семейства Araliaceae Juss", 2001, "Биологические исследования на Горнотаежной ...", "", "", "http://scholar.google.com/scholar?cites=11642775617378434158&as_sdt=2005&scioldt=0,5&hl=en&num=20", 30, 2015-08-20, "CITATION"
33. 1, "...", МД Зерова, ДР Каспарян, ВА Коляда...", "Аннотированный каталог насекомых Дальнего Востока России. Том I. Перепончатокрылые.-", 2012, "Влад.: ...", "msubiology.info", "", "http://scholar.google.com/scholar?cites=14709385965817510041&as_sdt=2005&scioldt=0,5&hl=en&num=20", 33, 2015-08-20, "PDF"
34. 1, "ВС Солдатов, АГ Воеводин, ВА Коляда", "Модель генерации поверхностных состояний в МДП-структурах при туннельной инъекции", 1990, "Поверхность", "", "", "http://scholar.google.com/scholar?cites=13410483039253831131&as_sdt=2005&scioldt=0,5&hl=en&num=20", 34, 2015-08-20, "CITATION"
35. 1, "ИЯ Боярский, АЮ Власов, МС Коляда", "О приборе для измерения поверхностной скорости водокаменных селевых потоков", 1970, "Вестник МГУ», сер. геогр", "", "", "http://scholar.google.com/scholar?cites=10286531389444088374&as_sdt=2005&scioldt=0,5&hl=en&num=20", 35, 2015-08-20, "CITATION"
36. 0, "...", АС Зеленко, НМ Иваньшина, МГ Коляда...", "Відомості про авторів", 2009, "ВІСНИК", "irbis-nbuv.gov.ua", "http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/vluf_2014_3_27.pdf", "http://scholar.google.com/scholar?cluster=4167166549794625333&hl=en&num=20&as_sdt=0,5", 36, 2015-08-20, ""
37. 0, "ВД Фролов, АС Коляда", "Рецензенты: к. б. н., старший научный сотрудник Горнотаежной станции ДВО РАН ТА Москалюк", 0, "", "", "", "http://scholar.google.com/scholar?q=related:mRwPBL-pZ9MJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5", 37, 2015-08-20, "CITATION"
38. 0, "...", СГ Коваленко, ТІ Ковбаса, МО Колесник, АМ Коляда...", "Відомості про авторів", 2009, "ВІСНИК", "chnpu.edu.ua", "", "http://scholar.google.com/scholar?cluster=5489135150129149377&hl=en&num=20&as_sdt=0,5", 38, 2015-08-20, "PDF"
39. 0, "ВД Гогунський, АС Коляда...", "Наукометричні дані наукового видання «Управління розвитком складних систем»", 2015, "Управління розвитком ...", "journals.uran.ua", "http://journals.uran.ua/index.php/2219-5300/article/view/38536", "http://scholar.google.com/scholar?q=related:9-0kHq1DTKEJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5", 39, 2015-08-20, ""
40. 0, "АС Коляда", "О ПЕРИДЕРМАЛЬНЫХ ДРЕВЕСНЫХ РАСТЕНИЯХ ЮГА ПРИМОРЬЯ", 0, "botsad.ru", "", "", "http://scholar.google.com/scholar?q=related:iby2IB9LOg4J:scholar.google.com/&hl=en&num=20&as_sdt=0,5", 40, 2015-08-20, "PDF"
41. 0, "...", ВО Яковенко, АС Коляда", "Розвиток системної концепції наукометричних баз даних", 2015, "Управління розвитком ...", "journals.uran.ua", "http://journals.uran.ua/index.php/2219-5300/article/view/38418", "http://scholar.google.com/scholar?q=related:kh1wI0iKvYAJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5", 41, 2015-08-20, ""

42. 0,"...", АЛ Гордиенко, ТМ Мельник, ЕН Коляда..."Выписка из приказа о допуске к кандидатским экзаменам № 32-19 от 27.04. 2010 г.",2010,"Филологические науки",,,,,,42,2015-08-20,"CITATION"
43. 0,"АК КОЛЯДА, АМ ВАЙСЕРМАН, ДС КРАСНЕНКОВ..."Исследование длины теломер у пациентов с болезнью Паркинсона",2014,,,,,"epigenetics.com.ua",,,,,,"http://scholar.google.com/scholar?q=related:S7MHNhkKWd4J:scholar.google.com/&hl=en&num=20&as_sdt=0,5",43,2015-08-20,"PDF"
44. 0,"ЯА Гринченко, АС Коляда","Древесные растения Уссурийского городского округа",0,"ББК 28.0 (2Р55) Ж 67 Составитель","uss.dvfu.ru",,,,,,44,2015-08-20,"PDF"
45. 0,"ДС Красненков, АК Коляда, АМ Вайсерман","Длина теломер в клетках крови людей разного возраста в популяции Украины",0,,,,,"http://scholar.google.com/scholar?q=related:glJPwFiPk4J:scholar.google.com/&hl=en&num=20&as_sdt=0,5",45,2015-08-20,"CITATION"
46. 0,"АС Коляда, НА Коляда","КЛАССИФИКАЦИЯ ПЛОДОВ ДРЕВЕСНЫХ РАСТЕНИЙ ПРИМОРСКОГО КРАЯ",2009,"Вестник Оренбургского ...","cyberleninka.ru","http://cyberleninka.ru/article/n/klassifikatsiya-plodov-drevesnyh-rasteniy-primorskogo-kрая",,"http://scholar.google.com/scholar?q=related:9w8hQu-bV78J:scholar.google.com/&hl=en&num=20&as_sdt=0,5",46,2015-08-20,""
47. 0,"...", РВ Козовий, ВП Колодченко, АК Коляда..."Авторский указатель",0,,,,,47,2015-08-20,"CITATION"
48. 0,"АС Коляда, НА Коляда","О ПУТЯХ РАСПРОСТРАНЕНИЯ ДИАСПОР ДРЕВЕСНЫХ РАСТЕНИЙ ПРИМОРСКОГО КРАЯ",2008,"Вестник Красноярского государственного ...","elibrary.ru",,,,,,48,2015-08-20,"CITATION"
49. 0,"АС Коляда, ВД Гогунский","Extracting information from the semistructured web pages",2014,"Eastern-European Journal of ...","journals.uran.ua",,"http://journals.uran.ua/eejet/article/view/19496",,"http://scholar.google.com/scholar?q=related:IWxTVKt8KyYJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",49,2015-08-20,""
50. 0,"...", НК Коваленко, ВП Колодченко, АК Коляда..."Авторский указатель",0,,,,,50,2015-08-20,"CITATION"
51. 0,"АФ Коляда, КС Кожемякин","Циркуляция кинетической энергии в сателлитах зубчатых механизмов",2014,"Вестник двигателестроения","irbis-nbuv.gov.ua",,,,,,"http://scholar.google.com/scholar?q=related:fn9oY-9kruUJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",51,2015-08-20,"PDF"
52. 0,"АС Коляда, ЮН Глущенко","О происхождении и синонимах русских названий птиц Приморья",0,"Animals and plants of the Russian Far East",,,,,,"http://scholar.google.com/scholar?q=related:kQMqbb63oNEJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",52,2015-08-20,"CITATION"
53. 0,"...", КА Маловичко, ОВ Ежова, ОВ Клочко, МГ Коляда..."РОЗДІЛ І. ТЕОРІЯ ТА ІСТОРІЯ ОСВІТИ 5",0,"irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Npdntu_pps_2014_1(1)_2.pdf",,,,,,53,2015-08-20,""
54. 0,"АС Коляда, ВД Гогунский","Authenticity of authorship of scientific publications using latent semantic analysis",2014,"Eastern-European Journal of ...","journals.uran.ua",,"http://journals.uran.ua/eejet/article/view/23942",,"http://scholar.google.com/scholar?q=related:hDuGb9MuMaMJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",54,2015-08-20,""
55. 0,"ВД Гогунський, ВО Яковенко, АС Коляда","ОСОБЛИВОСТІ ЦИТУВАННЯ НАУКОВИХ ПУБЛІКАЦІЙ У ІНТЕРНЕТ-ПРОСТОРИ",2015,,,,,"storage.library.opu.ua",,,,,,"http://scholar.google.com/scholar?q=related:1vChflpXLqYJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",55,2015-08-20,"PDF"
56. 0,"АС Коляда, ВА Волобоев, ВД Гогунський","СУЧАСНІ ТЕНДЕНЦІЇ РОЗВИТКУ СИСТЕМ АВТОМАТИЗАЦІЇ ТЕХНОЛОГІЧНИХ ПРОЦЕСІВ ЗА НУКОМЕТРИЧНИМИ ДАНИМИ SCOPUS",2014,,,,,"storage.library.opu.ua",,,,,,"http://scholar.google.com/scholar?q=related:xU29hao-l3wJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",56,2015-08-20,"PDF"

57. 0,"АС Коляда","Морфология надземных частей древесных растений Дальнего Востока в зимний период",2009,"Животный и растительный мир Дальнего Востока. ...",,,,,,"http://scholar.google.com/scholar?q=related:0rGvhicVvbEJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",57,2015-08-20,"CITATION"
58. 0,"...",ГК Жолудев, ЮГ Жолудева, ВВ Кідалов, АІ Коляда...,"ВІДОМОСТІ ПРО АВТОРІВ НОМЕРА",0,,,,,,58,2015-08-20,"CITATION"
59. 0,"ОФ Коляда, АФ Коляда, ТА Сокол, ДС Прокопенко...","МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ПОВЕРХНОСТИ ЗУБЧАТОГО РАБОЧЕГО ОРГАНА ДЛЯ ОБРАБОТКИ ПОЧВЫ",0,,,,,,59,2015-08-20,"CITATION"
60. 0,"КИ Лапин, НС Яковлева, АЮ Коляда","ГИДРАВЛИЧЕСКИЕ ПОТЕРИ ДАВЛЕНИЯ В ПРОТИВОПОЖАРНОМ ТРУБОПРОВОДЕ",0,"I. Методы и средства ликвидации аварий",,,,,,"http://scholar.google.com/scholar?q=related:M5Llo4NViesJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",60,2015-08-20,"CITATION"
61. 0,"...",ОФ КОВПИК, ЕА КОРНИЛОВ, ЮЕ КОЛЯДА...,"ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ НЕЛИНЕЙНЫХ ЭФФЕКТОВ ПРИ ВЗАИМОДЕЙСТВИИ НЕРЕЛЯТИВИСТСКИХ И РЕЛЯТИВИСТСКИХ ...",0,"iaea.org",,,,,,"http://scholar.google.com/scholar?q=related:1E48pkG8tl4J:scholar.google.com/&hl=en&num=20&as_sdt=0,5",61,2015-08-20,"PDF"
62. 0,"АС Коляда","ЖИВОТНЫЙ И РАСТИТЕЛЬНЫЙ МИР ДАЛЬНЕГО ВОСТОКА",0,"ibss.febras.ru",,,,,,"http://scholar.google.com/scholar?q=related:-Q-gr_0X3VQJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",62,2015-08-20,"PDF"
63. 0,"ДС Красненков, АК Коляда, НГ Ахаладзе...","Длина теломер в клетках крови у жителей Киевской области разного возраста",2014,"Проблемы старения ...",,"irbis-nbuv.gov.ua",,,,,,"http://scholar.google.com/scholar?q=related:MGxetWNBqqMJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",63,2015-08-20,"PDF"
64. 0,"ВД Гогунський, АС Коляда, ВА Яковенко","ПРОЕКТУВАННЯ СИСТЕМИ МОНІТОРИНГУ ПУБЛІКАЦІЙ НАУКОВЦІВ В НАУКОМЕТРИЧНИХ БАЗАХ ДАНИХ",0,"... та молодих вчених (Херсон, 5-6 ...",,"ek.ucoz.ua",,,,,,"http://scholar.google.com/scholar?q=related:SfGxuuzxo1sJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",64,2015-08-20,"PDF"
65. 0,"АС Коляда, АМ Кузнецов","Никольск",1997,"Уссурийский: штрихи к портрету",,,,,,65,2015-08-20,"CITATION"
66. 0,"СА Громов, АА Коляда...","КЛИНИКО-ДИАГНОСТИЧЕСКИЕ ОСОБЕННОСТИ ЭПИЛЕПСИИ, ОТЯГОЩЕННОЙ СОПУТСТВУЮЩЕЙ (НЕАКТИВНОЙ) ОРГАНИЧЕСКОЙ ...",2012,"Эпилепсия и ...",,"cyberleninka.ru",,,,,,"http://scholar.google.com/scholar?q=related:HNaQwnVh_yYJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",66,2015-08-20,"HTML"
67. 0,"АС Коляда, НА Коляда","ПУТИ ЕСТЕСТВЕННОГО ВЕГЕТАТИВНОГО РАЗМНОЖЕНИЯ СЕМЕННЫХ РАСТЕНИЙ ДАЛЬНЕГО ВОСТОКА",2011,"Вестник Красноярского государственного ...",,"elibrary.ru",,,,,,67,2015-08-20,"CITATION"
68. 0,"...",ОВ Епифанов, МН Шеремета, ВИ Коляда...,"Ульянов Петр Лаврентьевич",1990,"Матем. сб",,"mathnet.ru",,"http://www.mathnet.ru/rus/person9126",,"http://scholar.google.com/scholar?cluster=11414213825447204078&hl=en&num=20&as_sdt=0,5",68,2015-08-20,""
69. 0,"ВД Фролов, АС Коляда","Рецензент: доктор. с.-х. наук, профессор ГВ Гуков",0,,,,,,,"http://scholar.google.com/scholar?q=related:hh8J0fKZzj0J:scholar.google.com/&hl=en&num=20&as_sdt=0,5",69,2015-08-20,"CITATION"
70. 0,"АК Коляда, АС Соседко, МА Чивликлий...","МУТАЦИИ ГЕНОВ LRRK2, GBA и SNCA, АССОЦИИРОВАННЫХ С БОЛЕЗНЬЮ ПАРКИНСОНА У ЖИТЕЛЕЙ УКРАИНЫ",0,"geront.kiev.ua",,,,,,,70,2015-08-20,"PDF"
71. 0,"АС Коляда, ТА Мулина","Таксономический состав папоротников Уссурийского городского округа",0,"ББК 28.0 (2Р55) Ж 67 Составитель",,"uss.dvfu.ru",,,,,,71,2015-08-20,"PDF"
72. 0,"АС Уварова, СП Коляда","ПРОДОВОЛЬЧА БЕЗПЕКА ЗА ДОПОМОГОЮ МИТНО-ТАРИФНОГО РЕГУЛЮВАННЯ",2013,"ББК 65.9 (4укр)-55 I 66",,"sophus.at.ua",,,,,,"http://scholar.google.com/scholar?q=related:-e6D3IkAFyUJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",72,2015-08-20,"PDF"

73. 0,"СД Артамонов, АС Коляда","Животный и растительный мир Дальнего Востока. Вып. 13. Сб.",2009,"", "", "", "http://scholar.google.com/scholar?q=related:6_QM3-L1rzkJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",73,2015-08-20,"CITATION"
74. 0,"АС Голодных, КС Коляда","НАЧИСЛЕНИЕ АМОРТИЗАЦИИ В БУХГАЛТЕРСКОМ И НАЛОГОВОМ УЧЕТЕ (НА ПРИМЕРЕ ООО «НАУЧНО-ТЕХНИЧЕСКИЙ КОМПЛЕКС «...»,0,"scienceforum.ru", "", "", "http://scholar.google.com/scholar?q=related:Lbeq5XdHmwIJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",74,2015-08-20,"PDF"
75. 0,"АС Коляда","Этимология видовых эпитетов сосудистых растений Дальнего Востока России. Сообщение 6",0,"ББК 28.0 (2Р55) Ж 67 Составитель АС Коляда Ж 67 ...", "uss.dvfu.ru", "", "", 75,2015-08-20,"PDF"
76. 0,"...", МІ Іншин, СС Карінський, ІЯ Кисельов, ТА Коляда...","ПРАВОВИЙ МЕХАНІЗМ РЕГУЛЮВАННЯ ЗВІЛЬНЕННЯ З ІНІЦІАТИВИ ПРАЦІВНИКА",2014,"", "apdp.in.ua", "", "http://scholar.google.com/scholar?q=related:RZlcChOibgJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",76,2015-08-20,"PDF"
77. 0,"В Ляшенко, С Коляда, І Бережнюк, А Войцещук...","МИТНЕ РЕГУЛЮВАННЯ: ПРАВОВА ОСНОВА ТА ЗМІСТ",0,"Праці Одинадцяті міжнародної ...", "", "", "http://scholar.google.com/scholar?q=related:2JNuDsWGIZkJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",77,2015-08-20,"CITATION"
78. 0,"АС Коляда","Древесные растения Приморья (морфология в зимний период, транспирация хвойных высокогорья, биология и возможности использования ...",2008,"", "[б. и.] Владивосток", "", "", 78,2015-08-20,"CITATION"
79. 0,"...", НМ Коляда, ОА Маланова, А Плаушинайтене...","Редакционная коллегия сборника",0,"", "", "", "http://scholar.google.com/scholar?q=related:WG_f9T8ZOn8J:scholar.google.com/&hl=en&num=20&as_sdt=0,5",79,2015-08-20,"CITATION"
80. 0,"АС Зубченко, АА Коляда, ПИ Давидчук, ВА Ермилов","КОРРОЗИОННАЯ СТОЙКОСТЬ ХРОМИСТЫХ И ХРОМОНИКЕЛЕВЫХ СТАЛЕЙ И ИХ СОЕДИНЕНИИ В РАСПЛАВЕ СЕРНИСТОГО НАТРИЯ",1971,"Автоматическая сварка: ...", "", "", "http://scholar.google.com/scholar?q=related:lk9i9x0O_a0J:scholar.google.com/&hl=en&num=20&as_sdt=0,5",80,2015-08-20,"CITATION"
81. 0,"...", ВА Гавриленко, КС В'ячеславович, ТА Коляда","Відомості про авторів",0,"irbis-nbuv.gov.ua", "", "http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/econpr_2014_3_14.pdf", "http://scholar.google.com/scholar?cluster=9769426254939348922&hl=en&num=20&as_sdt=0,5",81,2015-08-20,""
82. 0,"...", М Коляда, М Кадемія, С Кізім, О Будник, А Тимур...","ДИДАКТИКА, МЕТОДИКА І НАВЧАННЯ",0,"irbis-nbuv.gov.ua", "", "", "http://scholar.google.com/scholar?cluster=15879927800927729190&hl=en&num=20&as_sdt=0,5",82,2015-08-20,"PDF"
83. 0,"ВД Фролов, АС Коляда","ТИПЫ ПЛОДОВ ВЕСЕННИХ РАСТЕНИЙ ЮЖНОЙ ЧАСТИ ПРИМОРСКОГО КРАЯ",2007,"Бюллетень Ботанического сада института ...", "botsad.ru", "", "http://scholar.google.com/scholar?q=related:39H_Z59HtoJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",83,2015-08-20,"PDF"
84. 0,"ВО Єгошина, ОМ Коляда, СВ Бруснік...","Показники резистентності у хворих на пневмонію різної етіології",2010,"Аннали ...", "irbis-nbuv.gov.ua", "", "http://scholar.google.com/scholar?q=related:vQTeFBxrfsJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",84,2015-08-20,"PDF"
85. 0,"...", ЛА Майстренко, ОА Андреєва, МК Коляда...","ТЕХНОЛОГІЇ ТА ДИЗАЙН",0,"irbis-nbuv.gov.ua", "", "http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/td_2013_4_1.pdf", "", 85,2015-08-20,""
86. 0,"...", АА Берегова, НЛ Колганова, ЄВ Коляда","Сучасні методи оцінювання якості життя пацієнтів в медичній косметології",2012,"Дерматологія та ...", "irbis-nbuv.gov.ua", "", "", 86,2015-08-20,"PDF"
87. 0,"...", АМ Нестеренко, ОМ Коляда","Влияние микроорганизмов на развитие гиперпластических процессов в носоглотке",2011,"Аннали ...", "irbis-

- nbuv.gov.ua", "", "http://scholar.google.com/scholar?q=related:QrUPGFNy5SsJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",87,2015-08-20,"PDF"
88. 0,"ХВ Маркова, ТІ Коляда, ВО Єгошина...", "Особливості цитокінового профілю в динаміці запального процесу верхніх дихальних шляхів різної нозології", 2012, "Аннали ...", "irbis-nbuv.gov.ua", "", "http://scholar.google.com/scholar?q=related:FV3yGN3b0BwJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",88,2015-08-20,"PDF"
89. 0,"ТА Коляда, СВ Венедіктов", "Програма і робоча програма навчальної дисципліни «Трудове право», 2009, "", "eprints.kname.edu.ua", "", "http://scholar.google.com/scholar?q=related:8ZMJGZaeHAQJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",89,2015-08-20,"PDF"
90. 0,"...", ОН ТЕРЕХОВ, ВС КОЛЯДА", "Разработка стенда для динамической балансировки молотильных барабанов комбайнов фирмы Claas после ремонта", 2012, "ИЗВЕСТИЯ ...", "cyberleninka.ru", "", "http://scholar.google.com/scholar?cluster=10899210762316281749&hl=en&num=20&as_sdt=0,5",90,2015-08-20,"PDF"
91. 0,"...", ОМ Жаданова, ЄА Зеленов, ТО Кашпур, МГ Коляда", "Булах ИП Культурно-исторический подход в теориях развивающего обучения Вербовський ОВ Зміст виховання культури міжнаціональних ...", 0, "", "", "", "http://scholar.google.com/scholar?cluster=17364490071778052877&hl=en&num=20&as_sdt=0,5",91,2015-08-20,"CITATION"
92. 0,"НТ Берберова, ВП Осипова, МН Коляда...", "РУССКОГО ОСЕТРА ДОБАВЛЕНИЕМ (3, 5-ДИ-ТРЕТ-БУТИЛ-4-ГИДРОКСИФЕНИЛ) МЕТИЛЕНДИФОСФОНОВОЙ КИСЛОТЫ (57) Реферат: Изобретение ...", 2009, "", "freepatent.ru", "", "http://scholar.google.com/scholar?q=related:8LiqIlNrQ7EJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",92,2015-08-20,"PDF"
93. 0,"ВС Коляда, ВА Тищенко", "КОМПЛЕКСНАЯ РЕАБИЛИТАЦИЯ ИНВАЛИДОВ МЕТОДАМИ И СРЕДСТВАМИ ФИЗИЧЕСКОЙ КУЛЬТУРЫ", 0, "Здоровьесбережение как инновационный аспект ...", "rsvpu.ru", "", "http://scholar.google.com/scholar?q=related:zDJJg-n_QJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",93,2015-08-20,"PDF"
94. 0,"...", АІ Болховітінова, ОМ Коляда", "Влияние препарата галавіт на цитокіновий дисбаланс при опроміненні щурів лінії wistar на моделі інфекційного процесу", 2008, "Аннали ...", "irbis-nbuv.gov.ua", "", "http://scholar.google.com/scholar?q=related:IJb-K_FwoMgJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",94,2015-08-20,"PDF"
95. 0,"АГ Юрченко, СД Исаев, ГГ Коляда, СС Новиков", "ВЛИЯНИЕ СИММЕТРИИ ВТОРИЧНЫХ КОМПЛЕКСОВ НА СПЕКТРЫ ПМР 2-ПРОИЗВОДНЫХ АДАМАНТАНА В ПРИСУТСТВИИ ...", 1974, "", "Google Scholar", "", "http://scholar.google.com/scholar?q=related:ESlyLn8z8DAJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",95,2015-08-20,"HTML"
96. 0,"...", О Єгоров, А Кандиба, М Каленський, С Коляда...", "ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ", 0, "irbis-nbuv.gov.ua", "", "", "http://scholar.google.com/scholar?cluster=1796135210673478628&hl=en&num=20&as_sdt=0,5",96,2015-08-20,"RTF"
97. 0,"ЮЕ Коляда, ВИ Доценко, ОС Недзельский...", "ПЕРЕДВИЖНАЯ ЭЛЕКТРОФИЗИЧЕСКАЯ УСТАНОВКА ГИТ-20 ДЛЯ ЭЛЕКТРОИМПУЛЬСНОЙ ОБРАБОТКИ ШЛАМОВ МЕТАЛЛУРГИЧЕСКОГО ...", 0, "ELECTRICAL ...", "vestnik.kpi.kharkov.ua", "", "http://scholar.google.com/scholar?q=related:cj9CMXYTDaJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",97,2015-08-20,"PDF"
98. 0,"ТІ Коляда, СВ Бруснік...", "Оцінка рівня деяких цитокінів при пневмонії різноманітної етіології в експерименті", 2010, "Аннали ...", "irbis-nbuv.gov.ua", "", "", 98,2015-08-20,"PDF"
99. 0,"...", ВВ Глоба, КВ Маркова, ТІ Коляда", "Микробиологические и иммунологические показатели у больных с хроническими гиперпластическими заболеваниями верхних дыхательных путей", 2012, "Аннали ...", "irbis-nbuv.gov.ua", "", "http://scholar.google.com/scholar?q=related:uaNtQwbHPWMJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",99,2015-08-20,"PDF"
100. 0,"І Каромбович, С Коляда, В Ляшенко, А Мостовий...", "ПРОБЛЕМА ЗАСТОСУВАННЯ РЕЖИМУ МИТНОГО СКЛАДУ", 0, "", "", "", "http://scholar.google.com/scholar?cluster=11325907126713277372&hl=en&num=20&as_sdt=0,5",100,2015-08-20,"CITATION"

101. 0,"...", С Бричок, Т Глуцук, В Гончарук, А Горбатюк...,"Наші автори",0,"irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Ppps_2009_28_30.pdf",,"http://scholar.google.com/scholar?cluster=14611450807723799710&hl=en&num=20&as_sdt=0,5",101,2015-08-20,""
102. 0,"...", СО Качанов, АВ Кім, ОП Коляда...,"Відомості про авторів",2009,"Управління ...",,"irbis-nbuv.gov.ua",,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Uprv_2009_2.pdf&P21DBN=UJRN&Z21ID=#page=174",,"http://scholar.google.com/scholar?cluster=9001907418163333184&hl=en&num=20&as_sdt=0,5",102,2015-08-20,""
103. 0,"...", ОС Чирков, АВ Зорницький, ОБ Коляда...,"Semper tiro: Студентський науково-літературний часопис",2012,,,,,"eprints.zu.edu.ua",,"http://eprints.zu.edu.ua/7702/",,"http://scholar.google.com/scholar?q=related:in50TOv8fiYJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",103,2015-08-20,""
104. 0,"...", А Онищенко, С Ус, ВД Дербенцев, ЮВ Коляда...,"ВВ Вітлінський, ГП Суханова. Кількісне оцінювання",0,"irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Mise_2009_80_2.pdf",,,,,,104,2015-08-20,""
105. 0,"ЮТ Пименов, АИ Фоменко, СА Коляда","Введение в химию высокомолекулярных соединений",2003,,,,,"Астрахань",,,,,,105,2015-08-20,"CITATION"
106. 0,"АВ Гулевич, СГ Коляда","Оценки температурных и оптических неоднородностей в кювете лазера с ядерной накачкой",0,"Препринт","Обнинск",,,,,,106,2015-08-20,"CITATION"
107. 0,"...", СС Карінський, ІЯ Кисельов, ТА Коляда...,"ЮС Тихонович ПРАВОВИЙ МЕХАНІЗМ РЕГУЛЮВАННЯ ЗВІЛЬНЕННЯ З ІНІЦІАТИВИ ПРАЦІВНИКА",2014,"ББК 67.0 я43 А ...",,"apdp.in.ua",,,,,,"http://scholar.google.com/scholar?q=related:2XtAUvce-mQJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",107,2015-08-20,"PDF"
108. 0,"...", ТГ Куценко, ЮВ Белаш, ОЮ Прижкова, ТА Коляда...,"Методологические аспекты формализации основных направлений инноваций в области экономики, организации и управления в строительной отрасли ...",2006,,,,,"eprints.kname.edu.ua",,"http://eprints.kname.edu.ua/31642/1/26.pdf",,,,,,108,2015-08-20,""
109. 0,"...", АВ КОКАРЕВ, ГС МОВКАЛОВА, СГ КОЛЯДА...,"ОСОБЛИВОСТІ ДІАГНОСТИКИ ЦВС-2 ЗА ВИКОРИСТАННЯ МОЛЕКУЛЯРНИХ МЕТОДІВ ДОСЛІДЖЕННЯ",0,"biosafety-center.com",,,,,,"http://scholar.google.com/scholar?q=related:w1CJW8aHxwEJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",109,2015-08-20,"PDF"
110. 0,"...", ЛФ Каськова, П Якубова, ГІ Овчиннікова, ОО Коляда...,"ЗМІСТ CONTENTS Тези",0,"irbis-nbuv.gov.ua",,,,,,110,2015-08-20,"PDF"
111. 0,"...", АЮ Волянський, ОМ Коляда...,"Латентна інфекція як фактор ризику швидкого прогресування атеросклеротичного процесу",2006,"Аннали ...",,"irbis-nbuv.gov.ua",,,,,,"http://scholar.google.com/scholar?q=related:2A0QYB5CA8MJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",111,2015-08-20,"PDF"
112. 0,"...", ПІ Ігнатенко, ВІ Капланов, ЮЄ Коляда...,"Волошин В'ячеслав Степанович, д-р техн. наук, проф.",2008,,,,,"irbis-nbuv.gov.ua",,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/vpdy_2008_18(1)_1.pdf",,,,,,112,2015-08-20,""
113. 0,"ОН КОЛЯДА, СИ ПОХИЛ, АВ Тупотилов...","Методы верификации основного этиологического фактора экзогенных аллергических альвеолитов в г. Комсомольске",2013,"... проблем біології і ...",,"cyberleninka.ru",,,,,,"http://scholar.google.com/scholar?q=related:pnm9bEKfSfIJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",113,2015-08-20,"HTML"
114. 0,"...", ІК Бистряков, ТМ Паянок, ЛА Задорожня, ТА Коляда...,"ФІНАНСИ УКРАЇ",0,"irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Fu_2013_8_2.pdf",,"http://scholar.google.com/scholar?cluster=6913820157769107202&hl=en&num=20&as_sdt=0,5",114,2015-08-20,""
115. 0,"...", ЕА КОРНИЛОВ, ЮЕ КОЛЯДА...,"IAEA-CN-28/E-9 ЭКСПЕРИМЕНТАЛЬНЫЕ ИССЛЕДОВАНИЯ НЕЛИНЕЙНЫХ ЭФФЕКТОВ ПРИ ВЗАИМОДЕЙСТВИИ

- НЕРЕЛЯТИВИСТСКИХ И ...",1971,"Plasma ...","International Atomic Energy Agency",,,,,,"http://scholar.google.com/scholar?q=related:4-4wcSIYXgEJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",115,2015-08-20,"CITATION"
116. 0,"...", ВБ Бублик, СВ Коляда","ОБ АЛГОРИТМЕ ДЕСЕКВЕНЦИИ ЦИКЛИЧЕСКИХ ОПЕРАТОРОВ Киев В работе излагается алгоритм десеквенции циклических операторов, ...",1983,"Оптимизация ...","Академия наук СССР, Сибирское ...",,,,,,"http://scholar.google.com/scholar?q=related:GnladJ8q_BsJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",116,2015-08-20,"CITATION"
117. 0,"БН Ньюшков, ВС Пивцов, НА Коляда...","Стабилизация волоконного фемтосекундного лазера по оптическому стандарту частоты с использованием электрооптического кристалла КТР",2015,"Квантовая ...","quantum-electron.ru",,"http://www.quantum-electron.ru/php/paper_rus.phtml?journal_id=qe&paper_id=15768",,,,,,117,2015-08-20,""
118. 0,"...", Е Басистая, ЕВ Вторых, ИА Анпилогова, ЛИ Коляда...","Секретарь Стадник Наталья 24-96 ШЧГ Постольник Александр Яковлевич 22-42 52-42 ШЧЗ Кеодан Владимир Юрьевич 22-41 ШЧЗС Соснов ...",0,"Инженер",,,,,,118,2015-08-20,"CITATION"
119. 0,"...", ОМ Ярмач, ВЮ Докучаева, ДВ Радько, ОП Коляда","МЕТОДОЛОГІЧНІ ЗАСАДИ ДЕРЖАВНОСТІ, ПРАВА І БЕЗПЕКИ",0,"irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Pib_2009_5_2.pdf",,,,,,119,2015-08-20,""
120. 0,"...", ОВ Ігнатова, НВ Коваленко, ОП Коляда...","Відомості про авторів",2009,"Управління ...","irbis-nbuv.gov.ua",,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&Z21ID=&Image_file_name=PDF%2FUprv_2009_1.pdf&IMAGE_FILE_DOWNLOAD=1#page=181",,"http://scholar.google.com/scholar?cluster=4864297468364801937&hl=en&num=20&as_sdt=0,5",120,2015-08-20,""
121. 0,"...", АО Ігнатенко, ВП Капітон, ОВ Капітон, ІГ Коляда","Про авторів",0,"irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Fkzh_2013_39_31.pdf",,,,,,121,2015-08-20,""
122. 0,"СА Коляда","ПОДРАЖАНИЕ КАК ОДНА ИЗ ДЕТЕРМИНАНТ ПРЕСТУПНОСТИ НЕСОВЕРШЕННОЛЕТНИХ",2010,"Российский криминологический взгляд",,"elibrary.ru",,,,,,"http://scholar.google.com/scholar?cluster=242920350174293479&hl=en&num=20&as_sdt=0,5",122,2015-08-20,"CITATION"
123. 0,"НА Коляда, БН Ньюшков, АВ Иваненко...","Генерация диссипативных солитонов в сверхдлинном волоконном лазере с активной синхронизацией мод",2013,"Квантовая ...","mathnet.ru",,"http://www.mathnet.ru/rus/qe15041",,"http://scholar.google.com/scholar?cluster=9635290906304828931&hl=en&num=20&as_sdt=0,5",123,2015-08-20,""
124. 0,"...", К Неля, В Коваль, Н Коляда, А Коростелёв...","ВИКОРИСТАННЯ ІДЕЙ ЯНА АМОСА КОМЕНСЬКОГО У КОНТЕКСТІ ЗАГАЛЬНОСВІТОВИХ ТЕНДЕНЦІЙ МОДЕРНІЗАЦІЇ СЕРЕДНЬОЇ ТА ВИЩОЇ ...",0,"komeniana.udpu.org.ua",,,,,,"http://scholar.google.com/scholar?q=related:wIObgvF0MP0J:scholar.google.com/&hl=en&num=20&as_sdt=0,5",124,2015-08-20,"PDF"
125. 0,"...", А Кентій, В Лозицький, М Слободянюк, ІА Коляда...","Наукова література, видана Інститутом історії України НАНУ в 2010 р.",0,"irbis-nbuv.gov.ua",,,,,,"http://scholar.google.com/scholar?cluster=17475907739896424121&hl=en&num=20&as_sdt=0,5",125,2015-08-20,"PDF"
126. 0,"...", ВА Гавриленко, СВ Коверга, ТА Коляда...","Сведения об авторах",0,"irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/econpr_2014_3_15.pdf",,,,,,126,2015-08-20,""
127. 0,"...", НС ГОРБУНОВА, ОА КОЛЯДА","ФРАКЦИОННО-ГРУППОВОЙ СОСТАВ СОЕДИНЕНИЙ MN В ЧЕРНОЗЕМАХ СЕГРЕГАЦИОННЫХ КАМЕННОЙ СТЕПИ",2012,"... . Серия: Естественные и ...",,"cyberleninka.ru",,,,,,"http://scholar.google.com/scholar?q=related:iba_iKhof_MJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",127,2015-08-20,"HTML"

128. 0,"...", ЮГ Килина, ОВ Кожевина, НИ Коляда...,"НАШИ АВТОРЫ",2008,"Вестник Алтайского ...","asau.ru","http://www1.asau.ru/doc/nauka/vestnik/2008/5/author.pdf",,,,,128,2015-08-20,""
129. 0,"ЕС Коляда, ЮА Козлова","Система государственного регулирования деятельности предприятия",0,,,,,"http://scholar.google.com/scholar?q=related:tbCJXOv7nwJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",129,2015-08-20,"CITATION"
130. 0,"...", ГВ Хмельницкий, СВ Данчук, ЕЛ Коляда,"Послеоперационная ликворея у больных с височнобазальными опухолями головного мозга",2007,"УКРАЇНСЬКИЙ ...","cyberleninka.ru",,,,,,"http://scholar.google.com/scholar?q=related:jJhuovH_gcsJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",130,2015-08-20,"HTML"
131. 0,"ОА Журан, КОС Интернет-маркетинг","ИНТЕРНЕТ-МАРКЕТИНГ КАК ИНСТРУМЕНТ ДЛЯ ПРОДВИЖЕНИЯ САЙТА",0,"economics.opu.ua",,,,,,"http://scholar.google.com/scholar?q=related:1LPmkCPtPPMJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",131,2015-08-20,"PDF"
132. 0,"ЛБ Кириллова, СА Коляда","СПРАВНИТЕЛЬНАЯ ХАРАКТЕРИСТИКА МЕТОДОВ ОПРЕДЕЛЕНИЯ А-ТОКОФЕРОЛА",1999,"Вестник Астраханского ГТУ. Химия и ...","Google Scholar",,,,,,"http://scholar.google.com/scholar?q=related:I0Q5qa5Ka9wJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",132,2015-08-20,"HTML"
133. 0,"...", А Пермінова, Н Побірченко, С Самаріна,"Наши авторы",0,"journals.uran.ua",,,,,,"http://journals.uran.ua/index.php/2307-4833/article/download/15426/13267","http://scholar.google.com/scholar?cluster=916217879769637028&hl=en&num=20&as_sdt=0,5",133,2015-08-20,""
134. 0,"...", АА Крапівіна, ОЮ Качур, ЮБ Яценко, П Коляда...,"ЗдоРов'я населення: тенденції та Прогнози",0,"irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/Uzn_2014_1_2.pdf",,,,,134,2015-08-20,""
135. 0,"...", НО Коллякова, ТЮ Колотова, ОК Коляда...,"Авторський покажчик",2013,"Журн. НАМН ...","irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/jnamnu_2013_19_4_33.pdf",,,,,135,2015-08-20,""
136. 0,"ТІ Коляда, ОА Романова, ТА Сидоренко...","Оцінка впливу низькоінтенсивного пренатального опромінення на гематологічний статус новонароджених щурів",2004,"Вісник наукових ...","irbis-nbuv.gov.ua",,,,,136,2015-08-20,"CITATION"
137. 0,"РИ Гарбер, ВС Карасев, ВМ Коляда...","РАДИАЦИОННОЕ РАСПЫЛЕНИЕ МОНО-И ПОЛИКРИСТАЛЛОВ МЕДИ В ПОЛЕ ИЗЛУЧЕНИЯ ЯДЕРНОГО РЕАКТОРА",1969,"Журнал ...","Академия Наук СССР",,,,,,"http://scholar.google.com/scholar?q=related:mrGrwflBD80J:scholar.google.com/&hl=en&num=20&as_sdt=0,5",137,2015-08-20,"CITATION"
138. 0,"...", Т Кияк, К Ліліана, А Приходько, К Еліна, А Кондрук...,"Аджабі Яссін, Застровський Олександр Семантичний рівень мовної особистості. Особливості інтерпретації слів арабською та німецькою мовними ...",0,"irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/akpif_2015_2_2.pdf",,,,,138,2015-08-20,""
139. 0,"...", ФВ Климовицкий, РВ Козовий, ОК Коляда...,"Авторский указатель",0,"irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/PSD_2013_22_4_13.pdf",,,,,139,2015-08-20,""
140. 0,"...", С Заброварний, С Каріков, Т Ковальчук, І Коляда...,"Коротко про авторів",0,"history.org.ua",,,,,,"http://history.org.ua/JournALL/green/11/51.pdf","http://scholar.google.com/scholar?cluster=8511315069036912745&hl=en&num=20&as_sdt=0,5",140,2015-08-20,""
141. 0,"КВ Коляда, СП Певцов, АИ Федоренко","Объекты логистической инфраструктуры нескольких поколений",2012,"Логистика сегодня","grebennikon.ru",,,,,,"http://scholar.google.com/scholar?cluster=2276666303662538467&hl=en&num=20&as_sdt=0,5",141,2015-08-20,"CITATION"
142. 0,"...", ОП Ковальчук, ВД Ковальчук, ЕК Коляда...,"ФОНЕТИЧНА, ЛЕКСИЧНА ТА ГРАМАТИЧНА СИСТЕМИ МОВИ ТА МЕТОДИ ЇХ ДОСЛІДЖЕНЬ",0,"irbis-nbuv.gov.ua",,,,,142,2015-08-20,"PDF"

143. 0,"...", ВС Кисличенко, СІ Климнюк, ОМ Клімова, ТІ Коляда...,"Заступник головного редактора: Похил Сергій",0,"irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/ami_2012_1_1.pdf",,,,,143,2015-08-20,""
144. 0,"...", Г Владимов, А Еременко, В Ерофеев, С Есин...,"Современный литературный процесс в России",0,"mgup-dki.narod.ru",,,,,,"http://scholar.google.com/scholar?q=related:b-0v6LfzmdMJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",144,2015-08-20,"DOC"
145. 0,"...", СВ Михайлова, ОВ Тупотілов, ОМ Коляда...,"Розділ 1. ЕКОЛОГІЧНА БЕЗПЕКА",0,"irbis-nbuv.gov.ua",,,,,,"http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/ebpk_2014_16_2.pdf",,,,,145,2015-08-20,""
146. 0,"...", Л Набок, І ДЕМУЗ, С ЧЕРНЕЦЬКА, І КОЛЯДА...,"РОЗДІЛ І. ПИТАННЯ ІСТОРІЇ УКРАЇНИ, ЗАГАЛЬНОЇ ІСТОРІЇ, ІСТОРІОГРАФІЇ ТА МЕТОДОЛОГІЇ",0,,,,,"http://scholar.google.com/scholar?cluster=9573764883202221725&hl=en&num=20&as_sdt=0,5",146,2015-08-20,"CITATION"
147. 0,"ТІ Коляда, СВ Бруснік...,"Фагоцитарна активність нейтрофілів периферійної крові та макрофагів бронхоальвеолярного лаважу при пневмоній різної етіології в експерименті",2010,"Аннали ...",,"irbis-nbuv.gov.ua",,,,,,"http://scholar.google.com/scholar?q=related:RQ-X8ssrl9UJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",147,2015-08-20,"PDF"
148. 0,"ЕЛ КОЛЯДА, АЯ ГЛАВАЦКИЙ...,"ИНТРАКРАНИАЛЬНЫЕ ПОСЛЕОПЕРАЦИОННЫЕ ГНОЙНО-ВОСПАЛИТЕЛЬНЫЕ ОСЛОЖНЕНИЯ У НЕЙРООНКОЛОГИЧЕСКИХ БОЛЬНЫХ",2006,"УКРАЇНСЬКИЙ ...",,"cyberleninka.ru",,,,,,"http://scholar.google.com/scholar?q=related:liDM84znZaoJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",148,2015-08-20,"HTML"
149. 0,"ДА Коляда, ТН Коровина, СВ Савинова","Комплексная программа по сохранению и укреплению здоровья младших школьников",2012,"Волгоград «Учитель",,,,,,"149,2015-08-20,"CITATION"
150. 0,"...", А Марков, ИН Сеницын, А Антонова, МН Коляда...,"ФИЗИЧЕСКАЯ ХИМИЯ LA rus TR Дисторсия октаэдров Мо⁶⁺ О₆ и параметры модели валентностей связи Мо-О в неорганических ...",2008,"ДОКЛАДЫ АКАДЕМИИ ...",,,,,,"150,2015-08-20,"CITATION"
151. 0,"ТИ Коляда, СВ Бруснік, ІД Андреева...,"Некоторые особенности иммунного ответа под влиянием различных доз ионизирующего облучения у животных и человека",2007,"Аннали ...",,"irbis-nbuv.gov.ua",,,,,,"http://scholar.google.com/scholar?q=related:G8MI9-AfBfwJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",151,2015-08-20,"PDF"
152. 0,"...", БН Ньюшков, ИИ Корель, НА Коляда...,"Разработка прототипа компактного волоконного синтезатора частот для мобильных фемтосекундных оптических часов",2014,"Квантовая ...",,"mathnet.ru",,"http://www.mathnet.ru/rus/qe15977",,"http://scholar.google.com/scholar?cluster=1671646321435432877&hl=en&num=20&as_sdt=0,5",152,2015-08-20,""
153. 0,"...", І Сахневич, С Анатолій, С Прокопчук, І Коляда...,"Філософський вимір",0,,,,,"http://scholar.google.com/scholar?cluster=11669989685827529321&hl=en&num=20&as_sdt=0,5",153,2015-08-20,"CITATION"
154. 0,"...", АА Кобозева, ОА Кобылин, МГ Коляда...,"ОБ АВТОРАХ",0,,,,,"154,2015-08-20,"CITATION"
155. 0,"АД Дюдюн, НС Коляда, ЛА Погребняк...,"ОПЫТ ЛЕЧЕНИЯ БАКТЕРИАЛЬНЫХ ИНФЕКЦИЙ КОЖИ ПРЕПАРАТОМ «ФЛОКСИУМ»",0,"vitapol.com.ua",,,,,,"http://scholar.google.com/scholar?q=related:FpPO_IsXL9YJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",155,2015-08-20,"PDF"
156. 0,"...", АЯ Вязовик, АВ Головченко, ТИ Коляда...,"Низкопоточная анестезия севофлюраном с обеспечением газообмена надгортанным воздуховодом I-gel",2014,"Медицина ...",,"irbis-nbuv.gov.ua",,,,,,"http://scholar.google.com/scholar?q=related:RR66_pk--DUJ:scholar.google.com/&hl=en&num=20&as_sdt=0,5",156,2015-08-20,"PDF"

ПРИЛОЖЕНИЕ Б
АКТ ВНЕДРЕНИЯ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ



УКРАЇНА

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
БУДІВНИЦТВА І АРХІТЕКТУРИ

03680, м. Київ-37, Повітрофлотський пр-т., 31, телефон: 241 5580

АКТ

*про впровадження результатів дисертаційних досліджень
аспіранта Одеського національного політехнічного університету Колядою А.С.
«Моделі і методи пошуку інформації у наукометричних базах даних»*

В рамках виконання комплексної науково-дослідної роботи Київським національним університетом будівництва і архітектури (КНУБА) та Одеським національним політехнічним університетом (ОНПУ) на тему «Методологічні основи створення інформаційного середовища управління науковими дослідженнями структурних одиниць ВНЗ МОН України» (Номер державної реєстрації НДР 0115U000330) впроваджено систему збору інформації із міжнародних наукометричних баз даних аспірантом ОНПУ Колядою А.С., яка використовується КНУБА в науковій діяльності університету, а саме при оцінці наукових результатів підрозділів університету. Аспірантом ОНПУ Колядою А.С. отримані наступні наукові результати:

– розроблено інформаційно-пошукову систему з автоматизацією процесів пошуку метаданих публікацій з різних наукометричних баз даних, яка включає програмні інструменти ідентифікації елементів контенту Веб-сторінок і імітує інтелектуальні дії користувачів, що дозволяє вилучати бібліографічні дані статей для визначення інтегральної оцінки публікаційної активності авторів наукових публікацій;

– розроблена в ОНПУ інформаційна технологія, яка в дисертації орієнтована на забезпечення особистих інформаційних потреб окремих науковців, що формалізована, як програмний інтерфейс (API – Application Programming Interface), включена фахівцями КНУБА в програмний комплекс для моніторингу групової публікаційної активності лабораторій, кафедр і університетів.

Вказані вище матеріали впроваджені у 2015 р. при виконанні НДР № 0115U000330. Досягнуто позитивний ефект щодо моніторингу публікаційної активності наукових груп з урахуванням публікацій, що індексовані в різних наукометричних базах даних.

Науковий керівник НДР № 0115U000330,
завідувач кафедри інформаційних технологій,
доктор технічних наук, професор



А.О. Білощицький

Проректор з наукової роботи та
міжнародних зв'язків,
доктор технічних наук, професор

В.О. Плоский