

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «САНКТ-ПЕТЕРБУРГСКИЙ
ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ПЕТРА ВЕЛИКОГО»

Институт компьютерных наук и технологий
Высшая школа интеллектуальных систем и суперкомпьютерных технологий

Дисциплина «Цифровые ресурсы в научных исследованиях»

ОТЧЕТ
по лабораторной работе №2
на тему
«Поиск источников информации в специализированных коллекциях»

Выполнил:

Студент группы 3540901/02001
Дроздов Н.Д.

«....» 2020г.,
(Подпись)

Проверил:

Бендерская Е.Н.

«....» 2020г.,
(Подпись)

Санкт-Петербург
2020 г.

Содержание

1	Лабораторная работа №2	2
1.1	Цель работы	2
1.2	Программа работы	2
1.3	Ход работы	4
1.4	Вывод	11

Лабораторная работа №2

1.1 Цель работы

Ознакомиться с различными источниками информации как среди российских источников, так и среди англоязычных источников. Овладеть навыками поиска в различных базах и различных типов публикаций (диссертация, ВКР, статья, книга и другие).

1.2 Программа работы

1. Определить цель поиска. Сформулировать критерии отбора информации
2. Выбрать для поиска одну специализированную базу (библиотеку или специализированную поисковую систему) научных статей на русском языке и две на английском языке из представленных ниже (выбор обосновать):

- <http://ufn.ru/ru/>
- http://www.mathnet.ru/index.phtml?option_lang=rus
- <https://www.elibrary.ru/querybox.asp?scope=newquery>
- <http://arxiv.org>
- <http://citeseerx.ist.psu.edu/>
- <https://scholar.google.com/>
- <https://academic.microsoft.com/home>
- <https://www.sciencedirect.com/>
- <https://link.springer.com/nature.com>

Результаты отбора источников представить в виде библиографического списка. Для систематизации набора источников использовать одну из программ для управления библиографической информацией (например, Mendeley).

3. Дополнить список из п.2. 1-2 научными ресурсами на русском и 1-2 научными ресурсами на английском языке и также провести информационный поиск. Результаты отбора источников представить в виде библиографического списка
4. Дополнить результаты информационного поиска работами, найденными в научной социальной сети <https://www.researchgate.net/> или <https://www.academia.edu/> Отметить насколько отличаются результаты поиска от п.2.

Оценить давность и общее число работ, найденных по п.2-4. Результат представить в виде гистограммы распределения работ по годам.

5. Найти защищенные в СПбГУ бакалаврские и магистерские работы близкие по теме к Вашей теме <https://search.rsl.ru/ru#ff=07.09.2020&s=fdatedesc>

Привести 1-2 наиболее релевантные. Оценить давность и общее число работ.

6. Найти защищенные в России кандидатские и докторские диссертации близкие по теме к Вашей теме <https://elib.spbstu.ru/> Привести 1-2 наиболее релевантные.

Оценить давность и общее число работ.

7. Проверить наличие запатентованных РИД (результатов интеллектуальной деятельности) по Вашей теме в системе ФИПС <https://new.fips.ru/iiss/> и в ресурсе, предоставляемом Яндекс <https://yandex.ru/patents> Поиск выполнить как по Российским, так и по Международным документам <https://new.fips.ru/about/vptb-otdelenie-vserossiyskaya-patentno-tekhnicheskayabiblioteka/patentnyy-p> php <https://new.fips.ru/elektronnye-servisy/internet-resursy/index.php>

Сравните результаты, оцените удобство поиска и релевантность полученных документов.

Оценить давность и общее число работ. Результат представить в виде гистограммы распределения работ по годам.

8. Поиск по базам выполненных или продолжающихся проектов - <https://www.rscf.ru/contests/search-project> и https://www.rfbr.ru/rffi/ru/project_search

Оценить давность и общее число проектов. Результат представить в виде гистограммы распределения проектов по годам.

1.3 Ход работы

Целью своего поиска ставлю изучить нынешнее положение в сфере интеллектуального анализа данных, основные методы. Критерии отбора информации:

- Актуальность. Все ресурсы должны быть старше 2010 года.
- Информативность. Информация должна быть значима для будущего аналитического отчета.
- Достоверность. Информация представленная в источнике должна быть достоверна.

Среди русскоязычных баз была выбрана eLibrary, по сколько она выдала больше результатов публикаций по запросу "распознавание образов чем остальные. Для сравнения:

- УФН - 2 результата
- Math-Net - 5 результатов
- eLibrary - 499744 результатов (Рисунок 1.1)

Научная электронная библиотека LIBRARY.RU

Поиск в библиотеке

Вход в библиотеку

IP-адрес компьютера: 188.243.29.37

Название организации: не определена

Имя пользователя:

Пароль:

Вход

Запомнить меня

Правила доступа

Регистрация

Забыли пароль?

Навигатор

- ЖУРНАЛЫ
- КНИГИ
- ПАТЕНТЫ
- ПОИСК
- АВТОРЫ
- ОРГАНИЗАЦИИ
- КЛЮЧЕВЫЕ СЛОВА
- РУБРИКАТОР

РЕЗУЛЬТАТЫ ПОИСКОВОГО ЗАПРОСА

ВСЕГО НАЙДЕНО ПУБЛИКАЦИЙ: 499744 из 35120225

№	Публикация	Цит.
1	ОСОБЕННОСТИ СОВРЕМЕННЫХ МЕТОДОВ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ Москвитин А.А., Сазиев Т.М. В сборнике: Современные методы интеллектуального анализа данных в экономических, гуманитарных и естественнонаучных исследованиях. Материалы международной научно-практической конференции. Российский экономический университет им. Г.В. Плеханова, филиал в г. Пятигорске Ставропольского края. 2016. С. 11-18.	3
2	НЕЙРОКОМПЬЮТИНГ В НОВОЙ НАУКЕ О СПОРТЕ Петрунин Ю.Ю. Нейрокомпьютеры: разработка, применение. 2013. № 8. С. 066-071.	3
3	ПОСТРОЕНИЕ МОДЕЛИ ПРОГНОЗИРОВАНИЯ РАЗВИТИЯ СЛОЖНЫХ СИСТЕМ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИИ ИНФОРМАЦИОННОГО МОДЕЛИРОВАНИЯ Бритков В.В., Булычев А.В. В сборнике: Системный анализ и информационные технологии (САИТ-2011). труды IV Международной конференции. 2011. С. 175-178.	0
4	ПОДХОД К ОБРАБОТКЕ И ОЦЕНКЕ РЕЗУЛЬТАТОВ РАБОТЫ ИПС НА ОСНОВЕ РАСПРЕДЕЛЕННОГО КОЛЛЕКТИВНОГО СБОРА И АНАЛИЗА ДАННЫХ Британишский О.Л. Естественные и технические науки. 2011. № 1 (51). С. 204-208.	0
5	ОБ ИСПОЛЬЗОВАНИИ МНОГОКОМПОНЕНТНЫХ ВИТАМИННО-МИНЕРАЛЬНЫХ КОМПЛЕКСОВ ДЛЯ ПРОФИЛАКТИКИ ЖЕЛЕЗОДЕФИЦИТНОЙ АНЕМИИ У БЕРЕМЕННЫХ Громова О.А., Торшин И.Ю., Тетрушвили Н.К., Гришина Т.Р., Гоголева И.В., Серов В.Н. Медицинский алфавит. 2018. Т. 2. № 13 (350). С. 6-19.	1
6	ДИФФЕРЕНЦИАЛЬНЫЙ ХЕМОРЕАКТОМНЫЙ АНАЛИЗ ГЛЮКОЗАМИНА СУЛЬФАТА И НЕСТЕРОИДНЫХ ПРОТИВОВОСПАЛИТЕЛЬНЫХ ПРЕПАРАТОВ: ПЕРСПЕКТИВНЫЕ СИНЕРГИЧНЫЕ КОМБИНАЦИИ Громова О.А., Торшин И.Ю., Лиля А.М., Наумов А.В., Рейер И.А., Каратеев А.Е. Современная ревматология. 2018. Т. 12. № 2. С. 36-43.	16
7	ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА САМОАНАЛИЗА КОСМИЧЕСКОГО КОРАБЛЯ Деревягин С.А. В книге: Гагаринские чтения - 2018. Сборник тезисов докладов XLIV Международной молодежной научной конференции. 2018. С. 115-116.	0
8	ВЫЯВЛЕНИЕ АНОМАЛЬНОГО ТРАФИКА ДИНАМИЧЕСКИХ ВЫЧИСЛИТЕЛЬНЫХ СЕТАХС МОБИЛЬНЫМИ ПОТРЕБИТЕЛЬСКИМИ УСТРОЙСТВАМИ	0

Возможные действия

- Следующая страница
- Выделить все публикации на этой странице
- Снять выделение
- Добавить выделенные статьи в подборку
- Новая подборка
- Добавить все страницы с результатами поиска в указанную выше подборку
- Вернуться к поисковой форме и изменить условия запроса
- Создать новый поисковый запрос
- Продолжить поиск среди найденных результатов

Рис. 1.1: Результаты поиска на портале eLibrary

А среди англоязычных баз были выбраны Google Scholar и Springer Link.

Пополнение библиографического списка с помощью eLibrary, Google Scholar и Springer Link

С помощью Google Scholar была найдена книга технологии анализа данных [1], которая может стать хорошим фундаментом для дальнейших поисков информации.

В базе eLibrary была найдена статья, посвященная особенностям современных методов искусственного анализа данных [2].

В базе Springer Link была найдена книга, посвященная Data mining [3].

Пополнение библиографического списка с помощью ResearchGate, Academia. Провести сравнительный анализ с предыдущими сервисами

На ResearchGate был найден источник [4].

Работы практически не отличаются, поскольку Google Scholar является агрегационным сервисом, который объединяет в себе в том числе и базы ResearchGate и Academia.

Диаграмма распределения работ по годам

Для построения был использован сервис Google Scholar, с помощью фильтрации за конкретный год (к примеру, 2012-2012) я получал общее количество публикаций за эту дату. По итогу получилась следующая диаграмма (Рисунок 1.2).

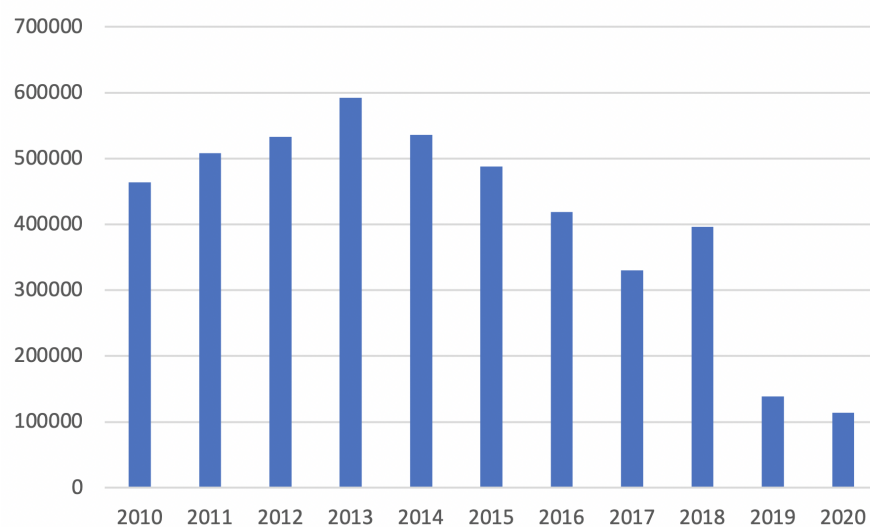



Рис. 1.2: Диаграмма распределения работ по годам с 2010

Основными причинами данного поведения диаграммы могут быть: по началу незнание или не интерес к теме, после чего сильное увлечение и анализ данной темы, после же снижение интереса к данной теме или большое количество закрытых (приватных) публикаций за последние годы, которые еще не были обнародованы.

Поиск близких бакалаврских и магистерских работ, защищенных в СПбПУ

По моему запросу было найдено 7066 квалификационных работ (Рисунок 1.3 студентов СПбПУ. Работы представлены 2013-2020 годов, что можно назвать достаточно актуальным. Наиболее релевантными мне показались работа бакалавра [5] и магистерская работа [6].

**ПОЛИТЕХ**
Электронная библиотека

Расширенный поиск | Атрибутный поиск | Контакты | Информационно-библиотечный комплекс

интеллектуальный анализ данных


С 17 марта 2020 г. для ресурсов (учебные, научные, материалы конференций, статьи из периодических изданий, авторефераты / особый режим использования. Обращаем внимание, что ВКР/НД

Поиск по электронным документам

Найдено документов: 11 006

Статистика слов: интеллектуальный: 70173, анализ: 1097847, данных: 3123302

1



[Методы и модели интеллектуального анализа данных в исследованиях в области медицины: бакалавра: направление 27.03.03 Системный анализ и управление ; образовательная программа математические методы системного анализа и управления в технических, экономических и с](#)

Авторы: Лукичева Дарья Михайловна

Организация: Санкт-Петербургский политехнический университет Петра Великого. Институт компьютер-

Выходные сведения: Санкт-Петербург, 2019

Коллекция: Выпускные квалификационные работы; Общая коллекция

Тематика: интеллектуальный анализ; модели интеллектуального анализа; методы интеллектуальног

Тип документа: Выпускная квалификационная работа бакалавра


Тип файла: PDF

Во второй части работы приведены практически полученные результаты обработки данных путем приме

модели, проведен анализ полученных результатов и освещена их практическая ценность.

pdf, 1,8 Мб

2



[Модели и алгоритмы интеллектуального анализа данных в сфере сервисного обслуживания работа бакалавра: 09.03.02 - Информационные системы и технологии](#)

Авторы: Витус Ольга Валерьевна

Организация: Санкт-Петербургский политехнический университет Петра Великого. Институт комп

Санкт-Петербург, 2018

<https://elilib.spbstu.ru/dl/3/2019/vr/vr19-3964.pdf/info>

Рис. 1.3: Результаты поиска в электронном каталоге СПбПУ

Поиск близких кандидатских и докторских диссертаций, защищенных в России

Было найдено более 23236 диссертаций (Рисунок 1.4) по запросу "интеллектуальный анализ данных" с датой публикации после 2015 года. В качестве наиболее релевантной была выбрана диссертация о коллективных методах интеллектуального анализа данных на основе нечеткой логики [7].

The screenshot shows the website of the Russian State Library (Российская государственная библиотека). The search bar at the top contains the query "интеллектуальный анализ данных". The results page shows 23,236 results. On the left, there is a sidebar with filters for document types, including "Диссертации" (Theses) which is checked. The main content area displays a list of results. The first result is a thesis by Цымблер, Михаил Леонидович, titled "Интеллектуальный анализ данных в СУБД". The abstract of the thesis is visible, discussing data mining and fuzzy logic.

Российская государственная библиотека

рус
eng

NEW Поступления А-Я Алфавитный указатель ? Справка

интеллектуальный анализ данных

Поиск

Расширенный поиск

↓ Единый электронный каталог INFO X

☐ Авторефераты диссертаций

☐ Газеты

☒ Диссертации 23236

☐ Изоматериалы

☐ Карты

☐ Книги (изданные с 1831 г. по настоящее время)

☐ Коллекция Шнейерсон

☐ Микроформы

☐ Ноты

☐ Рукописи

☐ Сериальные издания (кроме газет)

☐ Стандарты

Изданы с 2015 X Диссертации X очистить все фильтры

Результатов: 23 236, страница 1 / 2324 (0.532 сек.)

Внешние ресурсы: 2300

1

Цымблер, Михаил Леонидович.

Интеллектуальный анализ данных в СУБД : диссертация ... доктора физико-математических наук : 05.13.11 / Цымблер Михаил Леонидович; [Место защиты: Юж.-Ур. гос. ун-т (нац. исслед. ун-т)]. - Челябинск, 2019. - 260 с. : ил. [ещё](#)

[Цитаты из текста:](#)

стр. 14 интеллектуального анализа данных. Дается обзор публикаций, наиболее близко относящихся к теме диссертации. 1.1. Задачи интеллектуального анализа данных Под интеллектуальным анализом данных (Data Mining) понимают совокупность алгоритмов, методов и программного обеспечения для обнаружения в данных ранее неизвестных,

стр. 34 DMQL приведен на рис. 1.1. Язык DMQL позднее послужил основой для разработки целого ряда языков запросов интеллектуального анализа данных: язык для анализа временных данных TQML [61] Чена (X. Chen), 1998 г.; язык для анализа географических данных GMQL [100] Хана (Han), 1997 г.;

Рис. 1.4: Результаты поиска в электронном каталоге электронной российской библиотеки

Проверка запатентованных результатов интеллектуальной деятельности

Наибольшее количество результатов по запросы было в Яндексe, для сравнения:

- Яндекс - 10348 (Рисунок 1.5)
- ФИПС - 142 (Рисунок 1.6)

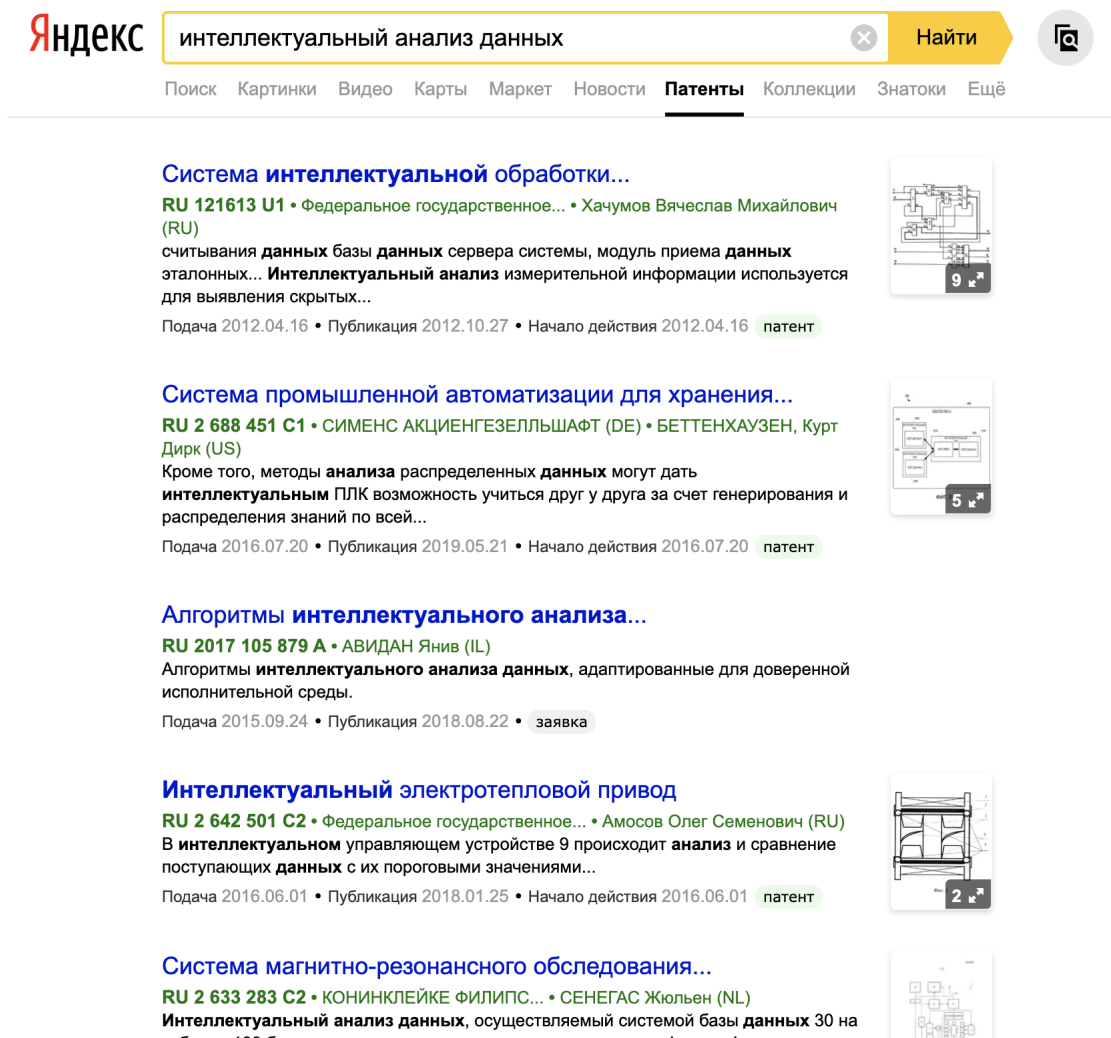


Рис. 1.5: Результаты поиска патентов в Яндексe

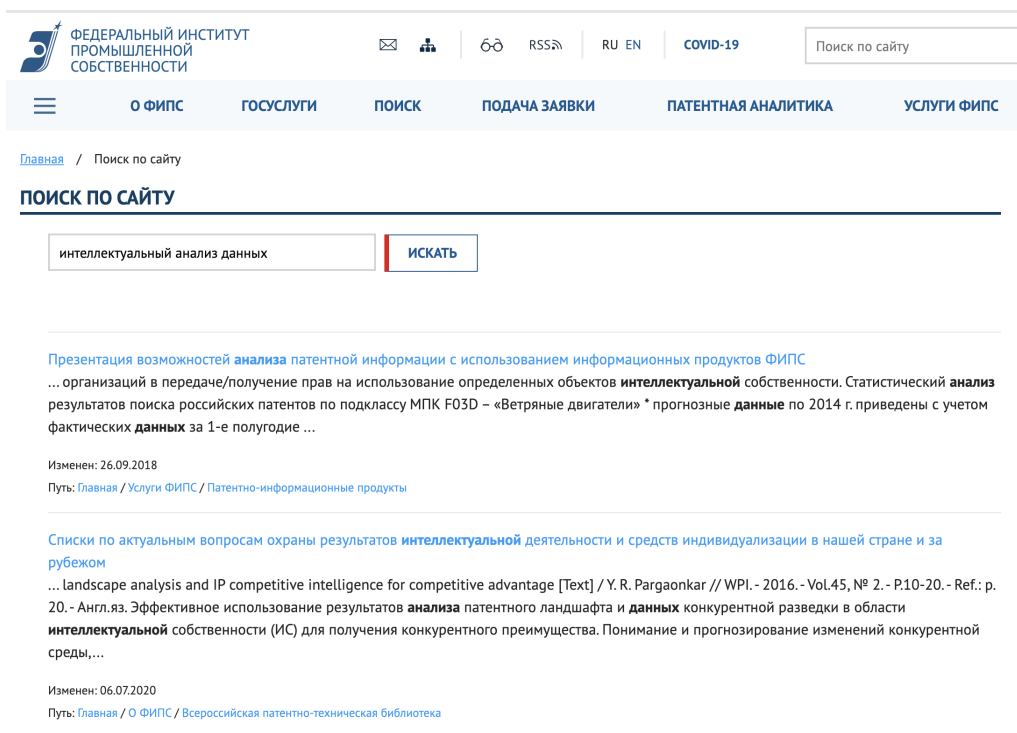


Рис. 1.6: Результаты поиска патентов в ФИПС


Яндекс имеет удобный пользовательский интерфейс, что облегчает знакомство с платформой и оставляет положительное впечатление для дальнейшего использования. Про платформу ФИПС такое сказать сложно. Диаграмма распределения патентов по годам (Рисунок 1.7 имеет нормальное распределение, что говорит скорее о спадающем интересе к моей теме.



Рис. 1.7: Диаграмма распределения патентов по годам в Яндекс

Поиск по базам выполненных или продолжающихся проектов

Поиск по базе российского фонда фундаментальных исследований не дал ни одного результата (Рисунок 1.8), а поиск по базе российского научного фонда выдал 18 результатов (Рисунок 1.9). Исходя из найденных результатов, можно сказать, что большинство проектов релевантны для моей темы, но их количество слишком мало для того, чтобы строить диаграмму распределения.



РОССИЙСКИЙ
ФОНД
ФУНДАМЕНТАЛЬНЫХ
ИССЛЕДОВАНИЙ

РФФИ

Карта сайта

Рос

Рус

Контакты

Обращения граждан

Противодействие коррупции

О ФОНДЕ

ПРЕСС-ЦЕНТР

КОНКУРСЫ

БИБЛИОТЕКА

ПОДАЧА ЗАЯВОК И СЕРВИСЫ

» Конкурсы » Поиск по проектам и заявкам »

Поиск по проектам и заявкам с 2012 года по настоящее время (пройдите по [ссылке](#))

Поиск по проектам и заявкам с 1998 по 2011 год (воспользуйтесь формой поиска ниже)

интеллектуальный анализ данных

ОБЛАСТЬ ПОИСКА

Везде

Поддержанные проекты

ТИП КОНКУРСА:

Все

ОБЛАСТЬ ЗНАНИЯ:

Все


ГОД ПРОВЕДЕНИЯ:

Все

Найти

По заданным критериям ничего не найдено

Рис. 1.8: Поиск в базе российского фонда фундаментальных исследований



Российский
научный
фонд

Конкурсы

О Фонде

Новости

Документы

Контакты

ИАС

ENG

Найти проекты

№	Номер проекта	Название, руководитель	Коды классификатора	Конкурс	Организация, регион
1	17-71-10176	Параллельные вычисления и методы целочисленного программирования для задач кластеризации и интеллектуального анализа больших массивов данных, <i>Ушаков А. В.</i>	основной: 01-202 дополнительные: 01-203, 01-214	Конкурс 2017 года по мероприятию «Проведение инициативных исследований молодыми учеными» Президентской программы исследовательских проектов, реализуемых ведущими учеными, в том числе молодыми учеными	ИДСТУ СО РАН, Иркутская обл
2	18-11-00087	Комплексный анализ акустических и электромагнитных сигналов для оценки уровня сейсмической опасности и предупреждения природных и техногенных катастроф, <i>Маратупец Ю. В.</i>	основной: 01-213 дополнительные: 01-219,	Конкурс 2018 года «Проведение фундаментальных научных исследований и поисковых научных исследований отдельными научными группами»	ИКИР ДВО РАН, Камчатский край
3	17-41-00002	Методы кластеризации и поиска в сетях большого размера, <i>Пардалос П.</i>	основной: 01-202 дополнительные: 01-203, 01-210	Конкурс на продление сроков выполнения проектов, поддержанных грантами Российского научного фонда по приоритетному направлению деятельности Российского научного фонда «Проведение	Национальный исследовательский университет "Высшая школа экономики"

Рис. 1.9: Поиск в базе Российского научного фонда

1.4 Вывод

В ходе данной работы были изучены различные базы информации, а так же получено оценочное представление о их содержании и цели поиска в них. Был приобретен навык поиска в различных сервисах, как российских, так и иностранных.

А так же составлен библиографический список источников для написания аналитического отчета по теме "Современные методы распознавания образов".

Литература

- [1] Барсегян А. Технологии анализа данных: Data Mining, Text Mining, Visual Mining, OLAP. 2 изд. – БХВ-Петербург, 2008.
- [2] Москвитин А. А., Созиев Т. М. Особенности современных методов интеллектуального анализа данных //Современные методы интеллектуального анализа данных в экономических, гуманитарных и естественнонаучных исследованиях. – 2016. – С. 11-18.
- [3] Aggarwal C. C. Data mining: the textbook. – Springer, 2015.
- [4] Chowdhary, Prof. (2020). Data Mining. 10.1007/978-81-322-3972-7_17.
- [5] Витус, Ольга Валерьевна. Модели и алгоритмы интеллектуального анализа данных в сфере сервисного обслуживания техники [Электронный ресурс]: выпускная квалификационная работа бакалавра: 09.03.02 - Информационные системы и технологии / О. В. Витус; Санкт-Петербургский политехнический университет Петра Великого, Институт компьютерных наук и технологий ; науч. рук. С. А. Нестеров. — Электрон. текстовые дан. (1 файл : 1,81 Мб). — Санкт-Петербург, 2018.
- [6] Проценко, Елизавета Геннадьевна. Методика повышения лояльности пользователей облачных сервисов на основании результатов интеллектуального анализа данных [Электронный ресурс] = Methodology of increasing user loyalty based on data mining results: выпускная квалификационная работа магистра: 27.04.03 - Системный анализ и управление ; 27.04.03_01 - Теория и математические методы системного анализа и управления в технических и экономических системах / Е. Г. Проценко; Санкт-Петербургский политехнический университет Петра Великого, Институт компьютерных наук и технологий ; науч. рук. С. А. Нестеров. — Электрон. текстовые дан. (1 файл : 2,2 Мб). — Санкт-Петербург, 2019.
- [7] Полякова, Анастасия Сергеевна. Коллективные методы интеллектуального анализа данных на основе нечеткой логики : автореферат дис. ... кандидата технических наук : 05.13.01 / Полякова Анастасия Сергеевна; [Место защиты: ФГБОУ ВО Сибирский государственный университет науки и технологий имени академика М.Ф. Решетнева]. - Красноярск, 2019. - 20 с. Системный анализ, управление и обработка информации OD 9 19-5/1462