

***Систематизация
информационных ресурсов в
электронной библиотеке на
основе информационно-
поискового тезауруса***

магистрант: *Байдилдаева Айзада*

Научный руководитель: д. ф.-м. н., профессор *Федотов А.М.*

В настоящий момент значительная часть информационных ресурсов хотя и переведена в цифровую форму, но недоступна широкому кругу научной общественности, а ресурсы, представленные в Интернет, разрознены, недостаточно систематизированы и структурированы. Поэтому, в процессе научно-педагогической деятельности очень часто необходима систематизация и классификация имеющихся информационных ресурсов.

Целью работы является создание системы классификации информационных ресурсов посвященных созданию и развитию электронной библиотеки на основе информационно-поискового тезауруса.

Достижение **поставленной цели** осуществляется путем решения следующих **задач**:

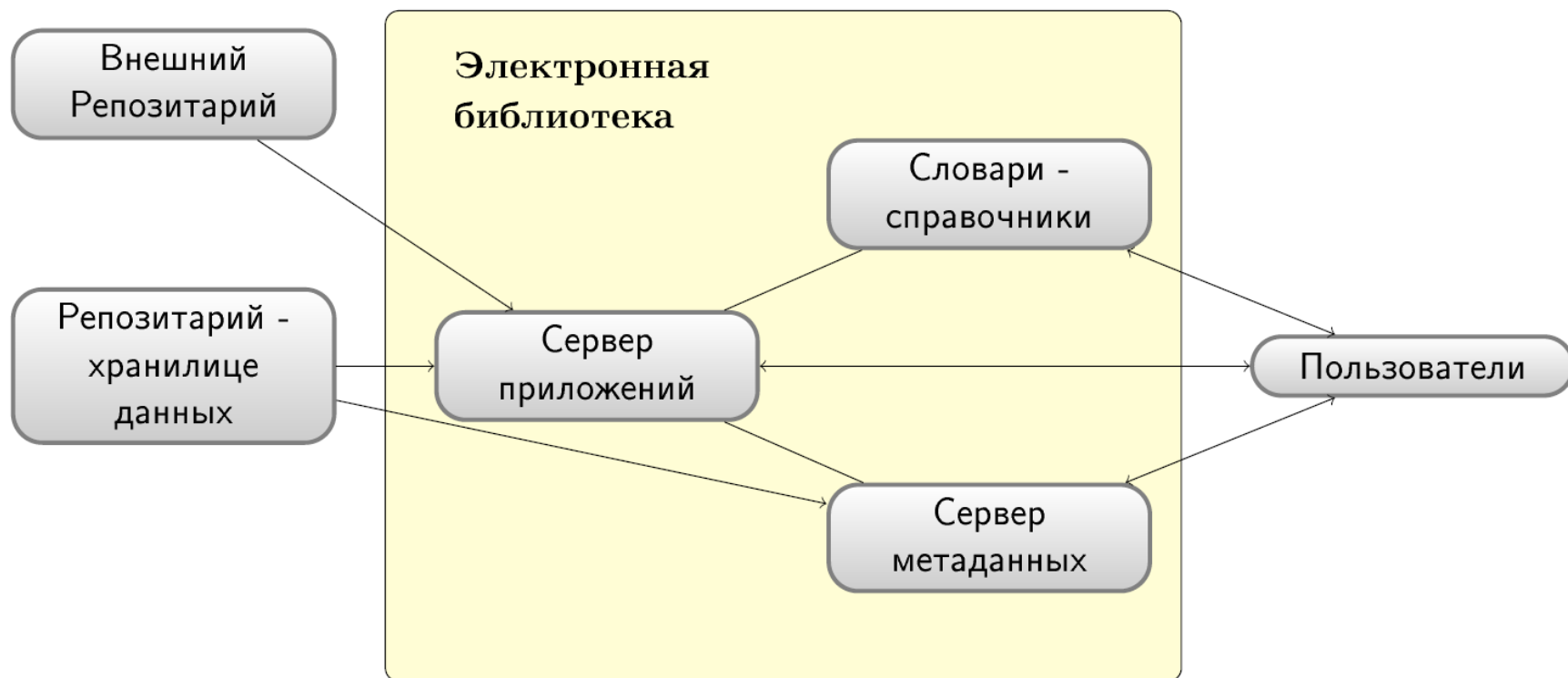
- 1) Сделать аналитический обзор международных стандартов и рекомендаций по организации баз данных публикаций, информационных ресурсов и электронных библиотек:
 - Изучить основные современные концептуальные модели электронной библиотеки;
- 2) Исследовать технологические средства для построения электронной библиотеки в области информационных систем;
- 3) Решить следующие вопросы для построения информационно-поискового тезауруса:
 - выявить классификационные признаки документов(публикаций) и систематизировать их;
 - определить и установить связи между документами;
 - расклассифицировать документы;
 - составить информационно-поисковый тезаурус и создать базу данных ключевых терминов.

Электронная (цифровая) библиотека — структурированная каталогизированная коллекция разнородных электронных документов, снабженная средствами навигации и поиска (в отличие от печатных изданий, микрофильмов и других носителей).

Функциональные требования к модели электронной библиотеки

- ✗ актуальность, полнота, достоверность происхождения документов;
- ✗ наличие большого числа словарей-классификаторов (справочников), для обеспечения идентификации и классификации ресурсов;
- ✗ поддержка неоднородных и слабо структурированных информационных ресурсов;
- ✗ поддержка взаимосвязей информационных ресурсов;
- ✗ предоставление информации пользователю в виде, выбранном пользователем;
- ✗ наличие программных интерфейсов для поддержки аналитической работы пользователя с помощью программных приложений;
- ✗ поддержка требований интероперабельности как на программном, так и на семантическом уровне;
- ✗ поддержка работы с внешними источниками.

Архитектура электронной библиотеки:



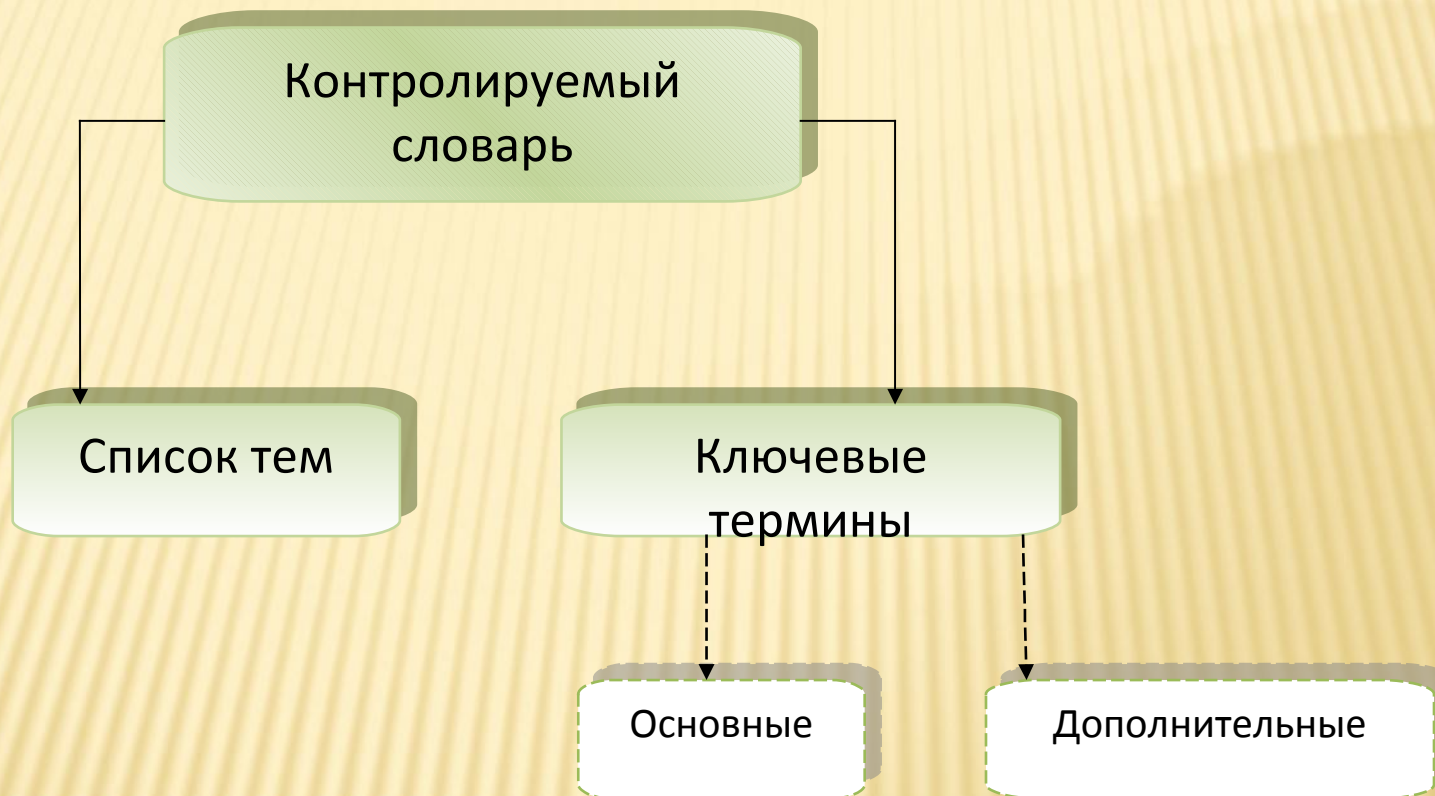
Систематизация электронных публикаций

Систематизация публикаций предназначена для их разделения по темам с целью сокращения времени поиска по запросам и выполняется с использованием ИПТ. Массивы публикаций сопровождаются метаданными, в состав которых обязательно включаются заголовки публикаций и списки ключевых слов. Процесс систематизации с целью сокращения затрат времени на его выполнение разделяется на два этапа: предметизацию и индексацию.

Информационно-поисковый тезаурус (ИПТ) — это контролируемый словарь терминов на естественном языке, явно указывающий отношения между терминами и предназначенный для информационного поиска.

Основными целями разработки традиционных ИИТ являются следующие:

- ✘ обеспечение перевода естественного языка документов и пользователей на контролируемый словарь, применяемый для индексирования и поиска;
- ✘ обеспечение последовательного использования единиц индексирования;
- ✘ описание отношений между терминами;
- ✘ использование как поискового средства при поиске документов;
- ✘ единицы традиционных информационно-поисковых тезаурусов.



Контролируемый словарь состоит из следующих элементов:

Персоны (persons_cat);

Ресурсы в коллекции (publ_cat);

Список разделов (тем) (temas_class).

ИИТ как система управления ключевыми терминами:

Корневые ключевые термины:

- ✘ Электронная библиотека
- ✘ Метаданные
- ✘ Каталогизация
- ✘ Репозиторий
- ✘ Архив
- ✘ Информационный поиск

Задачи при разработке ИПТ:

Установление связей между:

- ✘ ключевой термин → общие термины;
- ✘ ключевой термин → публикация;
- ✘ ключевой термин → персона.

Сибирское отделение РАН

БД "Электронная библиотека"



Страница администратора **Байдилдаева Айзада**

Редактирование документов в Коллекции publ_cat

Всего документов: 547

Редактирование документов

[1] [16] [31] [46] [61] [76] [91] [106] [121] [136] [151] [166] [181] [196] [211] [226] [241] [256] [271] [286] [301] [316] [331] [346] [361] [376] [391] [406] [421] [436] [451] [466] [481] [496] [511] [526] [541]

● [Розинг Борис Львович \(биография\)](#)

[show](#) [edit](#) [link](#) [delete](#)

● [Understanding Metadata // NISO Press. - 2004. 20 с.](#)

[show](#) [edit](#) [link](#) [delete](#)

● [Memex and Beyond \(Memex и его окружение\) Web-сайт \[http://cs.brown.edu/memex\]](#)

[show](#) [edit](#) [link](#) [delete](#)

● [Больман Александр-Наполеон Казимирович \(очерк о жизни и творчестве\)](#)

[show](#) [edit](#) [link](#) [delete](#)

● [Счеты русские \(Рассказ о русских счетах\)](#)

[show](#) [edit](#) [link](#) [delete](#)

● [Жизнь и деятельность В.Я. Вуняковского](#)

[show](#) [edit](#) [link](#) [delete](#)

● [Планиметр \(история создания\)](#)

[show](#) [edit](#) [link](#) [delete](#)

Ключевые термины, связанные с термином "Информационно-поисковый тезаурус":

1. [Вспомогательный дескриптор](#)
2. [Графический указатель информационно-поискового тезауруса](#)
3. [Дескриптор](#)
4. [Дескрипторный информационно-поисковый язык](#)
5. [Дескрипторный словарь](#)
6. [Иерархический указатель информационно-поискового тезауруса](#)
7. [Ключевое слово](#)
8. [Ключевые слова в контексте](#)
9. [Код дескриптора](#)
10. [Лексико-семантический указатель](#)
11. [Макротезаурус](#)
12. [Микротезаурус](#)
13. [Многоязычный информационно-поисковый тезаурус](#)
14. [Нондескриптор](#)
15. [Одноязычный информационно-поисковый тезаурус](#)
16. [Пермутационный указатель информационно-поискового тезауруса](#)
17. [Политематический информационно-поисковый тезаурус](#)
18. [Систематический указатель информационно-поискового тезауруса](#)
19. [Специализированный информационно-поисковый тезаурус](#)
20. [Язык ключевых слов](#)
21. [Контролируемое индексирование](#)

Ключевые слова: [тезаурус](#); [информационный поиск](#); [Автоматизированный информационный поиск](#);

● **Контекстный поиск:** Задайте образец для поиска:

Найти

Пример: Установленная связь между корневым термином «Информационно-поисковый тезаурус» и общими терминами.

Гипертекст [hypertext]

Гипертекст - принцип организации информационных массивов, при котором отдельные информационные элементы связаны между собой ассоциативными взаимосвязанными данными, взаимосвязанными указанными отношениями.

Термин гипертекст предложен в 1967 году [Теодором Нельсоном](#), представившим гипертекстовую структуру в виде ориентированного графа, в вершинах Средством организации связей в гипертексте и ориентации в нем является навигатор.

Идея использования ассоциативных отношений между документами впервые была высказана [Полем Отле](#) и частично реализована [Эмануэлем Гольдбергом](#) 1939 году, в которой описывает работу гипотетической машины MEMEX:

[Описание MEMEX из статьи В.Буша "As We May Think"](#)

Необходимость реализации прозрачного поиска информации в 1941 году формализует [Хорхе Борхес](#) в своем эссе "[Вавилонская библиотека](#)", описывая

Ключевые термины, связанные с термином "гипертекст":

1. [Навигатор \[navigator\]](#)
2. [Гиперссылка \[hyperlink\]](#)
3. [Ассоциативная связь](#)
4. [Гипертекстовая база данных](#)
5. [Гипертекстовый документ](#)
6. [Протокол передачи гипертекста](#)
7. [Язык гипертекстовой разметки HTML](#)
8. [Микрофиша](#)

Ссылки на публикации:

1. [Черняк Л. Статистическая машина Эмануэля Гольдберга // Открытые системы №03. 2004](#)
2. [Bush Vannevar. As We May Think // The Atlantic Monthly. July 1945.](#)
3. [Memex and Beyond \(Memex и его окружение\) Web-сайт \[http://cs.brown.edu/memex\]](#)
4. [Иванова И. А. Электронные гипертекстовые тезаурусы в дистанционном обучении/ Открытое образование. № 1. 2002. С.15- 19.](#)
5. [Борхес Х. Вавилонская библиотека // Хорхе Луис Борхес «Письмена Бога»: Республика. Москва. 1994.](#)

Ссылки на персон:

1. [Буш Вэннивер](#)
2. [Гольдберг Эмануэль](#)
3. [Нельсон Тед](#)
4. [Отле Поль](#)
5. [Энгельбарт Дуглас Карл](#)

Ключевые слова: поиск; [гипертекст](#); [навигатор](#); Статистическая машина;

Пример: 3 установленных связей: 1) между корневым термином «Гипертекст» и общими терминами; 2) между корневым термином «Гипертекст» и связанными с этим термином публикациями; 3) между корневым термином «Гипертекст» и персонами.

Словарная статья по термину «Метаданные»

Метаданные

Метаданные (Metadata) - данные о данных - данные, описывающие сущности, представленные в информационных системах, представляют собой характеристики описываемого информационного ресурса.

Метаданные - это содержание каталогов, справочников, реестры, и т.п., содержащие сведений о составе данных, содержании, статусе, происхождении, местонахождении

В настоящее время существует большое количество систем метаданных, предназначенных для описания различных классов информационных объектов. Использование с оригинальные системы метаданных и форматы метавписаний, а также разные подходы к решению прикладных задач.

Различаются описательные, структурные и административные метаданные.

Описательные метаданные описывают содержание информационного ресурса (например, это — набор значений элементов метаданных Дублинского ядра - Dublin Core), б

Структурные метаданные характеризуют общую структуру ресурса и ее компоненты, объем и другие подобные свойства описываемого ресурса.

Административные метаданные описывают даты создания и обновления ресурса, кем он создан или модифицирован, кто владелец прав на этот ресурс, полномочия дост

Стандарты метаданных:

Для обеспечения интероперабельности и повторного использования метаданных важное значение имеет стандартизация средств их представления. Действительность в этой с стандартов метаданных как независимых от сферы применения, так и предназначенных для специфических применений. К первой группе относятся, например, дескриптивные консорциума W3C: XML Schema, RDF, RDFS, OWL, OWL2 с его профилями и язык описания интерфейсов веб-сервисов WSDL; набор элементов Дублинского ядра (Dublin

Ключевые термины, связанные с термином "Метаданные":

1. [Дублинское ядро](#)
2. [EDIFACT](#)
3. [MATER](#)
4. [MARC](#)
5. [Стандарт для цифровых геопространственных данных \(SDGM\)](#)
6. [Формат обмена для справочников геопространственных данных \(DIF\)](#)
7. [Глобальная \(правительственная\) служба поиска информации\(GILS\)](#)
8. [Инициатива по кодированию текстов\(TEI\)](#)
9. [Шаблонно-ориентированные метаданные \(LAFA/WHOIS++\)](#)
10. [Кодировка архивных описаний\(ЕАД\)](#)

Ссылки на публикации:

1. [Understanding Metadata // NISO Press. - 2004. 20 с.](#)
2. [Ковалевский М.Р. Метаданные, их свойства, функции, классификация и средства представления // Труды 14-й Всероссийской научной конференции «Электронные Б](#)
3. [Лале Г.М., Самойлов А.В. Введение в электронные библиотеки // \[Электронный ресурс\] Москва, - 2005. - 92 с.](#)
4. [Иванченко И.Г. Подходы к созданию системы поддержки функционирования репозитория портала // Интернет-порталы: содержание и технологии: Об. науч. ст. Вып. 1](#)
5. [Бричковский В.И., Корсаков А.Г., Гурский Д.М. Технологии использования метаданных при создании электронных библиотек // Менеджмент библиотек учреждений об](#)
6. [Журбенко С.В., Лисовский К.А., Телицкий Г.С., Телицкий Э.С. Функциональные методы обработки слабоструктурированных данных и их применение для построения об](#)
7. [Наумова Е.А. Универсальные схемы метаданных и задача описания веб-ресурсов // Библиосфера. - 2006. - № 1. - с. 43-54.](#)
8. [Васильев А., Колесов Д., Сапегин С., Шамкина О. Извлечение метаданных и библиографических ссылок из текстов русскоязычных научных статей. // Труды конф](#)
9. [Борхес Х. Вавилонская библиотека // Хорхе Луис Борхес «Письмена Бога»: Республика. Москва. 1994.](#)

Ключевые слова: [Метаданные](#),

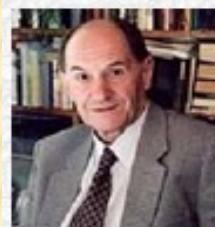
● [Комплексный поиск](#): Задайте образец для поиска:

[Найти](#)

Персоны в коллекции "Современные проблемы информатики"

Руджеро Сергеевич Гиляревский

Родился в Москве 31 августа 1929 г.



Основные достижения:

Гиляревский, Руджеро Сергеевич, доктор филологических наук, профессор, заслуженный деятель науки России, ведущий специалист в области социальной информатики, вёл в Этторе Матти и балерина Казанского оперного театра Екатерина Владимировна Крылаева. После отъезда отца из страны в 1929 г. его судьба неизвестна, мать в 1937 – 1954 гг. – у него женой, школьной учительницей Зинаидой Владимировной Гиляревской.

Краткая биография:

В 1947 – 1948 гг. Руджеро Сергеевич учился в Московском энергетическом институте на факультете электровакуумной промышленности, затем в МГУ им. М. В. Ломоносова на защитил кандидатскую диссертацию "Описание книг на иностранных языках для каталогов советских библиотек". 1989 г. – докторская диссертация "Общие закономерности в ра

Круг научных интересов Р.С. Гиляревского очень широк. Начаты еще в ВГБИЛ вместе с покойным Б.А. Старостиним работы по транскрипции иностранных имен привели к созданию справочника «Имена «Практическая транскрипция фамильно-именных групп» (2004, 2006). Книга «Основы научной информации», написанная в соавторстве с А.И. Михайловым и А.И. Черным, издана на семи языках в Берлине

В сферу научных интересов Р.С. Гиляревского входят проблемы научной информации и коммуникации, компьютерной технологии и, в частности, гипертекста и электронных книг, информационного менеджмента

В настоящее время Руджеро Сергеевич Гиляревский заведует Отделом теоретических и прикладных проблем информатики, Отделением научных исследований по проблемам информатики БИНИТИ, проф. Высшая школа экономики, является главным редактором сборников «Научно-техническая информация», заместителем главного редактора журнала «Международный форум по информатике», входит в Гиляревский награжден несколькими медалями, в 1999 г. ему присвоено звание "Заслуженный деятель науки Российской Федерации"

Научные труды:

1. Гиляревский Р.С. Информационный менеджмент : управление информацией, знаниями, технологиями. – СПб. : Профессия, 2009.
2. Гиляревский Р.С. Развитие принципов книгоописания : крат. очерк. – 2-е изд., доп. – СПб. : Профессия, 2008. – 240 с.
3. Гиляревский Р.С. Рубрикатор как инструмент информационной навигации / Р.С.Гиляревский, А.В.Шалкин, В.Н.Белоусов. – СПб. : Профессия, 2008. – 352 с.
4. Справочник информационного работника / науч. ред. Р.С.Гиляревский, В.А.Минкина. — 2-е изд., перераб. и доп.— СПб. : Профессия, 2007. — 384 с.
5. Информатика как наука об информации : информ., технол., экон., социал. и орг. аспекты / под ред. Р.С.Гиляревского. — М. : ФАИР-ПРЕСС, 2006. — 592 с.
6. Гиляревский Р.С. Основы информатики : курс лекций / Р.С.Гиляревский. — М. : Экзамен, 2003. — 319 с.
7. Рынок информационных услуг и продуктов / И.И.Родионов, Р.С.Гиляревский, В.А.Цветкова, Г.З.Залаев. — М. : МК-Периодика, 2002. — 549 с.
8. Информационное пространство новых независимых государств / Ю.М.Арский, Р.С.Гиляревский, И.И.Родионов, В.А.Цветкова и др. — М. : БИНИТИ, 2000. — 200 с.
9. Инфосфера : информационные структуры, системы и процессы в науке и обществе / Ю.М.Арский, Р.С.Гиляревский, И.С.Туров, А.И.Черный. — М. : БИНИТИ, 1996. — 489 с.
10. Михайлов А.И. Научные коммуникации и информатика / А.И.Михайлов, А.И.Черный, Р.С.Гиляревский. — М. : Наука, 1976. — 435 с.
11. Михайлов А.И. Основы научной информации / А.И.Михайлов, А.И.Черный, Р.С.Гиляревский. — М. : Наука, 1965. - 655 с.
12. Михайлов А.И. Информатика - новое название теории научной информации / А.И.Михайлов, А.И.Черный, Р.С.Гиляревский // Научно-техническая информация. - 1966. - № 12.

Публикации о жизни и деятельности персоны:

1. Черный, Л. Неожиданная информатика, или *what's be read* // Открытые системы №03, 2004

Публикации персоны:

1. [Гиляревский Р. С. Основы информатики : курс лекций — М. : Экзамен, 2003. — 319 с.](#)

Ключевые слова: [Информатика](#), [Библиотковедение](#).

🔍 **Лексиконский поиск:** Задайте образец для поиска:

[Найти](#)

Результаты:

В процессе работы были решены следующие задачи:

- Сделан аналитический обзор международных стандартов и рекомендаций по организации баз данных публикаций, информационных ресурсов и электронных библиотек;
- Исследованы технологические средства для построения ЭБ в области ИС;
- Изучены более 200 статей и монографий, посвященных организации ИС связанных с ЭБ, которые были занесены в хранилище данных DSpace, для части было сделано мета-описание, то есть информация была занесена в сервер метаданных и на основе обработки информации построен информационно-поисковый тезаурус посвященный тематике «Электронные библиотеки» и «Информатика»;
- Создан контролируемый словарь ключевых терминов, определены связи и расклассифицированы документы.

- Выступление с докладом на 51-й Международной Конференции МНСК с темой “Технология представления информационно-поискового тезауруса для электронной библиотеки по ИС”;

- Байдилдаева А.Б., Ручка Е.В, Федотов А.М., Федотова О.А. «Модель информационной системы для поддержки научно-педагогической деятельности.» // Вестник Новосибирского государственного университета. Серия: Информационные технологии. 2013. Т. 11. Вып. 3.

Благодарю за внимание!