

# НАУЧНО • ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

## Серия 1. ОРГАНИЗАЦИЯ И МЕТОДИКА ИНФОРМАЦИОННОЙ РАБОТЫ

ЕЖЕМЕСЯЧНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ СБОРНИК

Издается с 1961 г.

№ 4

Москва 2013

### МАТЕРИАЛЫ ВСЕРОССИЙСКОЙ НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ С ИНОСТРАННЫМ УЧАСТИЕМ «МЕТОДИКА ИСПОЛЬЗОВАНИЯ И ВЕДЕНИЯ УНИВЕРСАЛЬНЫХ КЛАССИФИКАЦИЙ НАУЧНОГО ЗНАНИЯ – УДК, ББК, ГРНТИ»

Москва, ВИНТИ РАН, ноябрь 2012 г.

26-29 ноября 2012 г. во Всероссийском институте научной и технической информации Российской академии наук при поддержке Российского фонда фундаментальных исследований (грант №12-07-06064) состоялась Всероссийская научно-практическая конференция «Методика использования и ведение универсальных классификаций научного знания – УДК, ББК, ГРНТИ», проведенная в рамках мероприятия в честь 60-летия ВИНТИ РАН. На конференции присутствовали представители научных и научно-технических библиотек, информационных центров, издательств и вузов из Алматы, Архангельска, Астаны, Белграда (Сербия), Владивостока, Калининграда, Москвы, Новосибирска, Рязани, Самары, Ухты, Хабаровска и других городов России.

Открыл конференцию директор ВИНТИ РАН, академик РАН **Ю.М. Арский**. Директор Научно-образовательного центра ВИНТИ РАН, заместитель директора по науке **И.Ю. Никольская** указала на важность участия библиотечных работников в научных работах по систематизации

документов и отметила, что проведение научно-практических конференций по УДК играет большую роль в повышении квалификации библиотечных и информационных работников.

В первый день пленарного заседания научно-практической конференции были заслушаны доклады о системах классификаций УДК и ГРНТИ, о современной структуре УДК. **В.Н. Белоозеров** сделал сообщение об универсальных системах классификации знаний в информационной практике. **И.Р. Смылова** и **И.М. Соловьева** сделали доклад о работах по классификационным системам, а **О.В. Кириллова** продемонстрировала возможности поиска информации в Тематических фрагментах Базы данных ВИНТИ РАН по ГРНТИ и УДК. **О.А. Антошкова** рассказала об издательской деятельности ВИНТИ РАН по УДК, представила обзор новых информационных продуктов по УДК, разработанных в ВИНТИ РАН. **Т.С. Астахова** осветила вопросы современной структуры УДК. **О.В. Смирнова** продемонстрировала методику составления индексов УДК, а

**В.Н. Белоозеров** привел сложные случаи индексирования по УДК. И закончился первый день работы научно-практической конференции докладом зав. библиотекой ИФП СО РАН **Н.Н. Шабуровой** о проблемах систематизации библиотечных фондов академических институтов Сибири.

Второй день пленарного заседания был посвящен опыту применения УДК в различных информационных и библиотечных системах. **Т.С. Маркарова** выступила с докладом о лингвистическом обеспечении систематизации и поиска информации в фондах научной педагогической библиотеки им. К.Д. Ушинского, а **А.Б. Антопольский** сделал сообщение о классификации информационных ресурсов. Об особенностях использования классификации УДК и ГРНТИ в ЦНСХБ Россельхозакадемии рассказала **Л.Н. Пирумова**. **М.К. Сухо руков** и **А.А. Гребенникова** осветили вопросы тематической рубрикации электронных изданий средствами УДК на примере научно-технического центра «Информрегистр» и затронули проблемы и перспективы, а также практику их применения. **А.А. Ивановский** поделился проблемами разработки тематического плана комплектования сети библиотек на основе УДК в Бен РАН.

В заключение второго дня выступил **Божидар Митрович** (Белград) с сообщением «УДК – способ предотвращения исследований истории славян до VI века или инструмент содействия организации национальных и международных междисциплинарных исследований истории славян до VI века».

Третье пленарное заседание было посвящено практическим вопросам организации классификационной деятельности в библиотечных фондах. **Т.С. Астахова** рассказала об изменениях в классах УДК 7/9 и привела примеры адаптации классификационных решений для книжного дела. **Э.Р. Сукиасян** прочел лекцию об алфавитно-предметном указателе, о технологии ведения АПУ в карточном и электронном форматах. Специали-

сты ВИНТИ РАН продемонстрировали новые возможности БД УДК и сделали презентацию нового информационного продукта «Эталонный мастер-файл УДК». К созданию этого продукта ВИНТИ РАН шел долгих 10 лет, с момента приобретения издательской лицензии УДК. **Н.Г. Сурикова** продемонстрировала все возможности мастер-файла для пользователей УДК. Демонстрация нового продукта вызвала оживленный интерес. Появился конкурент у БД УДК, который презентовали **Е.М. Зайцева** и **В.П. Анисимова** из ГПНТБ России. На этом заседании **Э.Р. Сукиасян** выступил с лекцией «Проблемы классификационной практики. Ситуационный анализ организационных и технологических решений», а специалисты ГПНТБ России сделали презентацию БД по ББК. Впервые в рамках конференции была широко освещена классификационная система ББК и продемонстрирована БД по ББК, разработанная ГПНТБ России.

По завершению работы пленарных заседаний было проведено очередное заседание методических советов по УДК и ГРНТИ.

Наиболее интересные материалы, из представленных на научно-практической конференции, публикуются в настоящем номере сборника «Научно-техническая информация».

При подведении итогов конференции участники единодушно поддержали предложение о ежегодном проведении таких методических мероприятий и отметили, что ВИНТИ РАН дает возможность получить новую информацию о развитии УДК, принять методическое решение при составлении индексов и повысить свою квалификацию.

Международный Консорциум УДК опубликовал информацию о научно-практической конференции на сайте и <http://www.udcc.org>.

По интересующим вопросам о методике совершенствования и применения УДК рекомендуем обращаться к авторам докладов по электронной почте: [typo@viniti.ru](mailto:typo@viniti.ru).

О.А. Антошкова (ВИНИТИ РАН)

## Универсальная десятичная классификация для библиотечного и книжного дела

*Рассматриваются вопросы применения Универсальной десятичной классификации в книжной отрасли. Приведены некоторые примеры совместимости книжной классификации КТК и УДК.*

**Ключевые слова:** книжное дело, универсальная десятичная классификация, книготорговая классификация (КТК)

Обмен данными в книжной отрасли требует использования стандартных средств тематической систематизации материала. На международном уровне одним из наиболее широко используемых средств систематизации является Универсальная десятичная классификация (УДК). В нашей стране по постановлению Правительства она является обязательной для индексирования всей научной и технической литературы с 1962 г., а с принятием новой версии ГОСТа на выходные сведения изданий УДК (наряду с ББК) стала необходима для оформления всей печатной продукции независимо от тематики. Согласно УДК систематизируются фонды основных российских научно-технических библиотек. Другие классификационные средства, такие как рубрикаторы информационных систем, международные классификации патентов, стандартов и т. п., имеют ограниченную сферу применения. Таким образом, УДК обладает потенциалом объединить книжные и другие информационные ресурсы единой сетью классификации, не зависящей от вида, носителя и источника документов. Иные универсальные международные системы классификации литературы, такие как Десятичная классификация Дьюи, Классификация Библиотеки Конгресса США, часто используемые в зарубежных информационных системах, не имеют традиции применения в нашей стране, что делает проблематичным их широкое внедрение в отечественную практику в обозримом будущем. Использование УДК прямо предусмотрено всеми форматами электронного обмена сведениями о книжной продукции - библиотечными форматами серии MARC, форматом обмена научно-технической информацией МЕКОФ и книготорговым форматом ONIX. Однако эффективное использование единой классификационной системы для упорядочения информационных потоков требует создания определённой инфраструктуры. Эта задача ложится на Всероссийский институт научной и технической информации Российской академии наук (ВИНИТИ РАН) как на головной орган Государственной системы научно-технической информации (ГСНТИ).

ВИНИТИ РАН - ведущий разработчик ряда классификационных систем и входит в Управляющий совет международного Консорциума УДК, владеющего правом интеллектуальной собственности на классификацию. В задачи ВИНИТИ РАН входит издание и продвижение в практику таблиц УДК на русском языке. Для достижения этой цели ВИНИТИ РАН реализовал разветвленную систему услуг и пособий, обеспечивающих эффективное использование УДК для тематической систематизации книжной продукции в нашей стране.

К настоящему времени завершено новое четвертое 10-томное издание эталонных таблиц УДК, насчитывающее около 500 авторских листов. Таблицы предоставляются пользователям как в печатном, так и в электронном виде. Они охватывают тематику всех отраслей знания и жанры художественной литературы. Следует обратить внимание на первый том (в котором представлена практика библиотечно-библиографического дела), на седьмой том (где содержится тематика типографской практики и книжной торговли) и на последний 10-й том (к которому относится большая часть материала - художественная литература, с которой работает книжное сообщество).

Регулярно издаются бюллетени изменений и дополнений. Сформирована база данных таблиц УДК на сервере ВИНИТИ РАН, которая позволяет производить анализ и сопоставление УДК с другими важными классификациями: Международной патентной классификацией, Государственным рубрикатором НТИ, российской Библиотечно-библиографической классификацией и др. Такое сопоставление обеспечивает совместимость УДК с систематическими каталогами отдельных информационных фондов, которые упорядочивают документы согласно особым классификациям, оптимальным в частных условиях. Например, классы Рубрикатора информационных изданий ВИНИТИ и Государственного рубрикатора НТИ снабжены ссылками на соответствующие классы УДК, что позволяет искать соответствующую литературу независимо от того, по какой системе она упорядочена.

На основе эталонного массива разрабатываются отраслевые рабочие таблицы, в частности для книжного дела. Представляется целесообразным гармонизировать с УДК российскую Книготорговую классификацию (КТК), рубрики которой были бы привязаны к соответствующим по содержанию классам УДК и допускали бы автоматическое перекодирование при заполнении полей в записях электронного обмена данными о книжной продукции. Предварительный анализ такой возможности свидетельствует о выполнимости и разумности постановки этой задачи.

При разработке КТК, предназначенной для унифицированной систематизации книжной продукции, с самого начала предусматривалось создание сопоставительных таблиц индексов КТК с индексами УДК.

Мы сможем предложить для издательского дела реализацию этого проекта в ближайшем будущем. Каждый уровень таблиц КТК будет снабжен соответствующими индексами УДК. Причем, во многих случаях возможно присвоение альтернативных вариантов индексов УДК рубрикам КТК для самостоятельного выбора издателями конкретного, соответствующего изданию, индекса.

Рассмотрим конкретный пример: Книга в жанре детектива

Рубрика КТК	Индекс УДК
Детектив	
64	82-312.4
Российский детектив	
640	82-312.4 или 821.161.1-312.4 (детектив, написанный на русском языке)
Зарубежный детектив	
641	82-312.4 (1-87) зарубежные детективы вообще или, например, 821.111.1-312.4 (детектив, написанный на английском языке)

В дальнейшем, предусматривая формирование единой автоматизированной базы данных на печатную продукцию и группирование литературы по определенной тематике, необходимо учитывать применение актуальной методики присвоения индексов УДК. Иначе создание единой базы данных весьма затруднено.

ВИНИТИ РАН ведет методическую работу по повышению квалификации пользователей УДК. Ежегодно проводятся методические семинары в рамках научно-практических конференций с привлечением слушателей всех регионов России и Ближнего Зарубежья. Материалы этих семинаров публикуются в журнале «Научно-техническая информация». Подготовлены и изданы учебные пособия, отражающие современное состояние информационно-практики. В Интернете ведется сайт, посвященный УДК. На сайте представлены сведения об изданиях УДК, проводятся интерактивные консультации для пользователей. На сайт можно послать вопрос о способе обозначения с помощью УДК той или иной тематики и оперативно получить ответ наших специалистов. Сборники этих ответов используются как пособия для обучения сложным случаям индексирования.

Для установления единства языка, на котором общаются разработчики и пользователи УДК, издан уникальный справочник по терминологии УДК на 5 языках (русский, английский, немецкий, испанский, французский).

Введены в действие Государственные стандарты по применению УДК, разработанные с активным участием ВИНИТИ РАН. Они составляют надежную нормативно-методическую основу применения УДК в библиотечном и книжном деле.

Основные стандарты системы СИБИД: ГОСТ Р 7.0.4 «СИБИД. Издания. Выходные сведения. Общие требования и правила оформления» – обязывает издателей проставлять индекс УДК в выходных данных всех книжных изданий.

ГОСТ 7.90 «СИБИД. Универсальная десятичная классификация. Структура, правила ведения и индексирования» – подробно описывает методику, структуру и функции УДК в информационной деятельности.

Примеры индексов, составленных по различным изданиям таблиц УДК

Индекс УДК, составленный по 3 <sup>МУ</sup> изданию	Индекс УДК, составленный по 4 <sup>МУ</sup> изданию
947 Литература по истории России	94(470) Литература по истории России
882 Художественная литература на русском языке	821.161.1 Художественная литература на русском языке

Применение УДК предусмотрено и другими стандартами: ГОСТ 7.5 «СИБИД. Журналы, сборники, информационные издания. Издательское оформление публикуемых материалов»; ГОСТ 7.19 «СИБИД. Формат для обмена библиографическими данными. Содержание записи»; ГОСТ 7.32 «СИБИД. Отчёт о научно-исследовательской работе. Структура и правила оформления»; ГОСТ 7.36 «СИБИД. Неопубликованный перевод. Общие требования и правила оформления»; ГОСТ 7.59 «СИБИД. Индексирование документов. Общие требования к систематизации и предметизации»;

ГОСТ 7.70 «СИБИД. Описание баз данных и машиночитаемых информационных массивов. Состав и обозначение характеристик»; ГОСТ 7.77 «СИБИД. Межгосударственный рубрикатор научно-технической информации. Структура, правила использования и ведения»; ГОСТ 7.83 «СИБИД. Электронные издания. Основные виды и выходные сведения»; ГОСТ Р 7.0.49 «СИБИД. Государственный рубрикатор научно-технической информации. Структура, правила использования и ведения» и др.

УДК 025.45.05

О.В. Смирнова (ВИНИТИ РАН)

## Составление индексов УДК. Методика и практика

*Приводится методика составления комбинированных индексов УДК на различные документы.*

**Ключевые слова:** *Универсальная десятичная классификация (УДК), индексирование документов, методика применения УДК, составление индексов УДК*

Универсальная десятичная классификация (УДК), созданная Полем Отле и Анри Лафонте-ном, существует более 100 лет. Она задумывалась как инструмент систематизации «Универсального библиографического репертуара» – карточного каталога литературы, изданной во всем мире по всем отраслям знания. В настоящее время УДК применяется во многих странах мира для систематизации печатной продукции, организации библиотечных каталогов и т.п.

Цель индексирования состоит в обеспечении средств навигации для поиска документов по заданной тематике и извлечения из них необходимых знаний. Одним из таких средств служит описание документа с помощью таблиц УДК. Результатом описания является классификационный индекс. В простых случаях для документа удается найти один индекс УДК, исчерпывающим образом описывающий его содержание. Такие индексы, включающие код одного класса Основной таблицы УДК, называют простыми. Современная тенденция индексирования состоит в том, чтобы в индексе документа отразить, по возможности, все классы, к тематике которых может быть отнесен документ в том или ином аспекте. Во многих случаях документ затрагивает несколько главных тем (отраслей знания), при этом в индексе требуется отразить также логическую и физическую форму документа. В этом случае индекс УДК складывается из кодов отдельных классов, представленных

в таблице наподобие фасетной формулы. Отсюда следует, что УДК не является перечислительной классификацией, а включает принципы координатного индексирования, аналитических и фасетных систем. Индексы, составленные из двух или нескольких классов Основной таблицы УДК, называют комбинированными. Как правило, они составляют систематизаторами в процессе индексирования документа. Результатом индексирования является поисковый образ документа как единый индекс, состоящий из последовательности цифр и других знаков, не разделенных пробелом.

Таблицы УДК, как известно, делятся на основные и вспомогательные. Основная таблица содержит понятия и соответствующие им индексы, с помощью которых систематизируется вся сумма знаний. В соответствии с десятичной системой вся таблица разделена на десять основных классов, каждый из которых в свою очередь подразделяется на десять более мелких подклассов и т.д.

Для наглядности и удобства чтения индекса после каждых трех его цифр, начиная слева, ставится точка.

Вспомогательные таблицы – это таблицы определителей, которые образуют отдельные классификационные ряды. Они служат для дальнейшей детализации индекса и отражают общие повторяющиеся для многих предметов признаки. Определители делятся на общие и специальные. Общие определители применяются во всех классах УДК,

специальные – только в определенном классе. Коды общих и специальных определителей последовательно присоединяют к индексу справа, после кодов основной таблицы. Как правило, определители присоединяют к основному индексу в следующем порядке:

- '1/9 специальные определители с апострофом
- .01/09 специальные определители с точкой ноль
- 1/-9 специальные определители с дефисом
- 02/-05 общие определители с дефисом
- (0...) общие определители формы документа
- (...) общие определители места
- "..." общие определители времени
- = общие определители языка
- (=...) общие определители народов.

Рассмотрим порядок индексирования документов. При индексировании документа можно придерживаться следующей схемы.

1. Определение предметного содержания документа, его основной тематики. Для этого находим в Основной таблице коды соответствующих классов.
2. Определение аспекта изложения темы. Для этого используются Таблицы общих определителей.
3. Обозначение формы представления документа, его языка и аспекта времени.
4. «Сборка» отдельных составных элементов (кодов) в классификационный индекс документа.

Содержательный анализ документа рассмотрим на следующем примере. Нужно поставить индекс УДК на статью "Тактико-технические характеристики океанского крейсера русского флота 80-х годов XIX века".

Во-первых, определяется предметное содержание документа. Главная его тематика – техника военно-морского флота. Находим в Основной таблице коды соответствующих классов:

623.8 Техника военно-морского флота. Военное кораблестроение. Боевые корабли и прочие суда

623.822.3 Крейсера.

Во-вторых, рассматривается аспект изложения темы. Это тактико-технические характеристики боевых кораблей. В классе 623 находим специальный определитель:

623-9 Тактико-технические характеристики и условия применения военной техники

В-третьих, обозначаются аспекты времени и места. В таблице общих определителей времени находим:

"18" XIX век

"188" Восемьдесятые годы XIX века.

В таблице общих определителей места находим: (470) Россия.

И, наконец, переходим к "сборке" составных частей индекса:

623.822.3-9(470)"188".

Другой пример: документам, посвященным Методическому семинару по УДК, прошедшему в Москве 27–29 ноября 2012 г., можно присвоить индекс

005.963.1:025.45.05УДК(470-25)"2012.11.27/.29", где:

005.963.1 Учеба. Тренинг. Учебные курсы

025.45 Десятичные (децимальные) классификации

025.45.05 Универсальные классификации (специальный определитель)

(470-25) Москва (общий определитель места)

"2012.11.27/.29" 27–29 ноября 2012 года (общий определитель времени).

В данном индексе использовано алфавитное расширение, чтобы подчеркнуть, что выбрана именно Универсальная десятичная классификация.

При систематизации нельзя ориентироваться только на название и ключевые слова документа. Необходимо ознакомиться с его содержанием. Например, для работы "Преемственность в стандартах оказания медицинской помощи пациентам с сердечно-сосудистыми заболеваниями", исходя из заголовка, можно присвоить индекс:

614.254.7:616.1-082, где:

616.1 Сердечно-сосудистые заболевания

616-082 Медицинское обслуживание. Организация лечения (специальный определитель)

614.254.7 Преемственность лечения при передаче больных другому врачу.

Однако речь в статье идет о переводе пациентов с одного стандарта лечения на другой. Следовательно, изменится и индекс УДК

616.1-082"713", где:

"713" Период преемственности. Результаты (общий определитель времени).

В заключение, необходимо напомнить о Рабочих таблицах УДК, в которых следует фиксировать все методические решения, принимаемые систематизаторами.

Т.С. Астахова, Е.Ю. Дмитриева, Н.Г. Сурикова (ВИНИТИ РАН)

## Эталонный массив УДК на русском языке

*Описывается новый вид информационного продукта, подготовленный специалистами ВИНТИ РАН на основе полной версии Таблиц УДК на русском языке для широкого распространения.*

**Ключевые слова:** УДК, эталонный массив, информационный поиск, релевантность, информационный продукт

Эталонный массив УДК сформирован на основе содержания десяти актуализированных томов 4-го полного издания Таблиц УДК на русском языке, в которые были в иерархической последовательности внесены все опубликованные ранее в отдельных шести выпусках изменения и дополнения. Кроме этого было проведено дополнительное редактирование всего информационного массива для устранения некорректного представления данных и более глубокого системного структурирования. Таким образом, на рынке информационных продуктов появляется эталонный объединенный массив, содержащий современные Таблицы УДК и включающий все изменения, дополнения и редакционные правки.

**Структура продукта.** Эталонный массив содержит все Основные Таблицы УДК, состоящие из 9 основных классов со специальными определителями внутри каждого класса и отдельно выделенные Вспомогательные Таблицы Iс/Iк, содержащие общие определители, которые выполняют функции стандартных классификационных фасет.

Сопоставимый с печатной версией Таблиц УДК эталонный массив отличается строгой упорядоченностью представления информации, глубоким структурированием по стандартным полям.

### Общая структура полей эталонного массива

- <01> иерархический код класса (индекс)
- <02> наименование класса
- <03><sup>1</sup> подзаголовок класса (резервное поле)
- <04> аспектная ссылка (отсылка)
- <05> методические указания
- <06> примеры комбинированных индексов
- <07> указание на подразделения по аналогии (подразделять как...)
- <08> примеры, относящиеся только к полю <07>
- <09> ссылки на другие классы
- <01И> иерархический код исключенного класса (индекс)
- <02> (наименование исключенного класса) Исключено [год]
- <09> ссылки на другие классы

<sup>1</sup> Это поле зарезервировано для обозначения подзаголовочных данных, которые до сих пор отсутствовали в русской версии и входили в поле <02>

- <опр01> иерархический код класса определителя (индекс)
- <опр02> наименование класса определителя
- <опр03> подзаголовок класса определителя (резервное поле)
- <опр04> аспектная ссылка (отсылка)
- <опр05> методические указания
- <опр06> примеры комбинированных индексов
- <опр07> указание на подразделения по аналогии (подразделять как...)
- <опр08> примеры, относящиеся только к полю <опр07>
- <опр09> ссылки на другие классы определителей

- <опр01И> иерархический код исключенного класса определителя (индекс)
- <опр02> (наименование исключенного класса определителя) Исключено [год]
- <опр09> ссылки на другие классы определителей.

За основу формирования эталонного массива была взята структура англоязычной стандартной (средней) версии УДК «Master reference file», предлагаемой своим пользователям международным Консорциумом УДК.

Преимущественным отличием русской версии является ее большая полнота (объем продукта превышает англоязычный аналог более чем в два раза), а также представление истории исключенных рубрик, т.е. наличие исключенных ранее классов и перекодировочных отсылок к новым классам.

**Поисковые возможности** нового продукта поддерживаются функциями операционной системы Microsoft Word. Это позволяет осуществлять поиск по всем возможным словоформам (по ключевым словам, с усечением, по отдельным фразам) в начале, в середине строки, а также по классам (индексам) с переходом на нужный класс.

Новый информационный продукт доступен всем пользователям, владеющим редактором Word.

Поисковые возможности эталонного массива нового информационного продукта поддерживаются функциями операционной системы Microsoft Word. Для поиска открывается диалоговое окно *Найти и заменить*. В командную строку диалогового окна в

зависимости от задачи, стоящей перед систематизатором, заносятся необходимые символы:

- а) цифровые – при поиске названия класса,
- б) буквенные – при составлении сложного индекса. При вводе полного класса (индекса) можно определить статус его актуальности на данный момент (действующий или исключенный).

#### *Варианты цифрового запроса*

Для осуществления поиска названия класса (индекса) вносится необходимый класс. Класс может быть как верхнего, так и нижнего уровня. При каждом нажатии мышки на экране последовательно выделяется запрашиваемый класс.

Сначала сделаем корректный запрос.

1. По цифровому коду, например, <01>**612.8**, где <01> – структурированное поле иерархического кода класса и находим название класса **Нервная система**.

2. По ключевому слову (фразе), например, фраза **электронный беспорядок**. В тексте находится и выделяется эта фраза, ей соответствует <опр01>**544.022.37** (запрос задается без скобок, так как название класса может включать несколько понятий и «электронный беспорядок» необязательно находится в начале строки). В данном случае получается, что при запросе одно структурированное поле – одно название класса (заданного уровня) и один запрос по ключевому слову – одно поле с иерархическим кодом класса, т.е. информационный шум сведен к нулю.

Теперь о поиске при некорректном запросе (как следствие, поиск длительный), т.е. алгоритм поиска «от общего к частному».

При вводе только индекса без метки поля, например **612.8**:

а) Word находит данное сочетание цифр в произвольном месте класса нижнего уровня (в середине, в конце, в начале), т.е. является его частью. В этом случае находится класс *необязательно верхнего уровня и при последовательном нажатии мышью выделенный фрагмент класса последовательно перемещается по строкам и проходит достаточно много времени пока не обнаружится класс заданного уровня;*

б) при этом найденный в результате поиска индекс может входить в любое из структурированных полей (в нашем случае <01>, <04>, <05>, <06>, <07>, <09>; а для определителей <опр01>, <опр04>, <опр05>, <опр06>, <опр07>, <опр09>).

*Результат:* визуально последовательно просматриваются все найденные варианты индексов, относящихся к разным полям.

Для сужения поиска добавляем к запрашиваемым цифрам элемент метки структурированного поля > (в нашем случае: >**612.8**):

а) в этом случае запрашиваемое сочетание цифр гарантированно будет находиться только в начале индекса нижнего уровня;

б) при этом найденный в результате поиска индекс также может входить в любое структурированное поле (в нашем случае <01>, <04>, <05>, <06>, <07>, <09>; а для определителей <опр01>, <опр04>, <опр05>, <опр06>, <опр07>, <опр09>).

*Результат* аналогичен предыдущему с единственной разницей: запрашиваемое сочетание цифр гаран-

тированно будет находиться только в начале индекса нижнего уровня, как и в предыдущем запросе визуально последовательно просматриваются все найденные варианты индексов, относящихся к разным полям, пока не обнаружится класс заданного уровня.

Для дальнейшего сужения поиска добавляется цифра метки структурированного поля, в нашем случае **1**, поскольку мы хотим получить название этого класса. Таким образом, мы предполагаем найти в массиве иерархический код класса 612.8 и соответственно – название этого класса.

В этом случае запрашиваемое сочетание цифр *гарантированно* будет находиться так же, как и в предыдущем запросе, только в начале класса нижнего уровня.

При этом найденный класс (индекс) в результате поиска будет иметь только одно структурированное поле (в нашем случае: <01>; для определителей – <опр01>). Первым находится класс высшего уровня, а далее – класс нижнего уровня.

*Результат* поиска соответствует полному корректному запросу <01>**612.8** и приводит к минимизации информационного шума.

Таким образом: 1) первым находится класс высшего уровня, а далее – класс более низкого уровня; 2) чем длиннее класс (индекс), тем меньше информационного шума и точнее соответствие поисковому образу (класс нижнего уровня находится сразу без продолжительного поиска); 3) мы получаем релевантный ответ на наш поисковый запрос.

#### *Варианты буквенного запроса.*

Операционная система Microsoft Word позволяет осуществлять поиск по всем возможным словоформам (с усечением ключевого слова, по ключевым словам, по отдельным фразам) в любом месте строки.

Поиск с усечением ключевого слова предполагает самый полный результат по найденным индексам. При дальнейшем увеличении количества буквенных символов происходит сужение поиска класса.

Нами был сделан корректный запрос – ввели целую фразу и получили код класса:

**электронный беспорядок** → <опр01>**544.022.37**.

Теперь сделаем запрос, который заведомо увеличит информационный шум, т.е. введем усеченную форму ключевого слова, например:

1) Запрос: **электрон**

Результат: электронный, электронная, электрон, электронов, фотоэлектронный и т.д.

2) Запрос: **электронный**

Результат: электронный разряд, электронный беспорядок, электронный перенос, гамма-электронный, электронный парамагнитный резонанс.

3) Запрос: **электронный беспорядок**

Результат: **электронный беспорядок** → **544.022.37**.

При поиске заданному названию класса найден соответствующий иерархический код класса.

Также возможно определить статус актуальности класса (индекса), например:

1) Запрос: **168.51**

Результат: <02>**Исключено как излишнее**

2) Запрос: **168.52**

Результат: <02>**Исключено** с указанием ссылок на другие классы (<09>165.321; <09>167.33).

Далее можно продемонстрировать поиск классов по ключевым словам для составления сложного индекса (см. с. 5-6).

Пример составления сложного индекса для названия статьи:

**Тактико-технические характеристики океанско-го крейсера русского флота 80-х годов XIX века.**

Ключевые слова:

- а) крейсер
- б) тактико-технические характеристики
- в) русский (российский) – усеченная форма – российский)
- г) XIX век
- д) 80-е годы.

При запросе «**крейсер**» выделяются несколько классов, учитывающих специфику данного термина. Здесь систематизатору нужно определиться, в каком аспекте употребляется это понятие. Выбираем класс **623.822.3**.

При запросе «**тактико-технические характеристики**» нас интересуют именно тактико-технические характеристики морского флота. Получаем ответ: **623-9**.

При запросе «**русский**» систематизатор должен определить употребление термина как «**российский**». Получаем: **(470)**.

Для уменьшения информационного шума при запросе «**XIX**» добавляем элемент метки >. Получаем ответ: **"18"**. Для обозначения десятилетий добавляем цифру 8 (по правилам составления индексов получаем: **"188"**). Далее составляется сложный индекс по принятым правилам и получаем требуемый сложный индекс: **623.822.3-9(470)"188"**.

**Возможное применение продукта.** Предлагаемый продукт поможет каждому индивидуальному или коллективному пользователю УДК, работающему с любым изданием или электронной версией УДК, сверить и откорректировать применяемые им классы (индексы) УДК, актуализировать свой информационно-поисковый аппарат. Это будет способствовать формированию единого информационного пространства в сегменте применения УДК как информационно-поискового языка. Кроме этого, жестко упорядоченная структура информационного массива может помочь начинающему пользователю освоить УДК как классификационную систему.

Глубокий уровень структурной разметки поможет пользователям самостоятельно формировать БД УДК для решения профессиональных задач.

УДК 025.2:025.45.05

А.А. Ивановский (Библиотека по естественным наукам РАН)

## **Применение УДК как основы лингвистического обеспечения сводного тематико-типологического плана комплектования ЦБС БЕН РАН: новые принципы**

*Автоматизированная система введения сводного тематико-типологического плана комплектования – одна из служебных информационно-поисковых систем Библиотеки по естественным наукам РАН. В её основе лежат две базы данных, связи между которыми строятся на основе индексов Универсальной десятичной классификации. Анализ принципов наполнения этих баз данных показал, что, с одной стороны, в индексировании близких тематик отсутствует воспроизводимость результатов такого индексирования, в том числе в силу его многоаспектности; с другой стороны, использование при индексировании проблемно-ориентированных, рабочих таблиц классификации делает невозможным использование при комплектовании индексов УДК, представленных на изданиях. Поэтому нами были выработаны новые принципы ведения сводного тематико-типологического плана комплектования, ключевым моментом которых является единая для всех участников системы лингвистическая основа – нормативная УДК. Это позволяет не только применять единые принципы определения формулировок и индексирования тем для всех научно-исследовательских учреждений РАН, но и использовать УДК как справочный аппарат для комплектователя по разным отраслям науки.*

**Ключевые слова:** Универсальная десятичная классификация (УДК), лингвистическое обеспечение, информационно-поисковая система, тематико-типологический план комплектования (ТТПК)

Библиотека по естественным наукам РАН (БЕН РАН) использует в своей работе разные классификационные системы. Для большинства задач применяется Универсальная десятичная классификация (УДК).

Универсальная десятичная классификация используется БЕН РАН в нескольких аспектах: для индексирования источников; для организации тематического поиска в электронном каталоге; в автоматизированной системе ведения сводного тематико-типологического плана комплектования БЕН РАН. Есть и другие сферы использования УДК в БЕН РАН.

Первые два из перечисленных аспектов отражают общеизвестное применение УДК, изложенное, например, в национальном стандарте ГОСТ 7.90–2007 [1].

Использование УДК как основы лингвистического обеспечения автоматизированной системы ведения сводного тематико-типологического плана комплектования (ТТПК) Централизованной библиотечной системы (ЦБС) БЕН РАН является примером расширения сферы применения этой классификационной системы.

Этапы формирования автоматизированной системы ведения сводного ТТПК ЦБС БЕН РАН изложены в ряде публикаций [2–5]. Изначально лингвистическое обеспечение сводного ТТПК строилось на основе Рубрикатора ГАСНТИ [6]. Затем лингвистической основой стала УДК. По мере развития автоматизированной системы роль УДК, как лингвистической основы ТТПК, несколько изменялась [7].

В связи с очередным пересмотром планов комплектования библиотек ЦБС БЕН РАН возникла возможность внести изменения в автоматизированную систему, не нарушая при этом хода технологических процессов.

Цель настоящей статьи – рассмотрение накопившихся проблем при использовании сложившейся системы использования УДК как лингвистической основы ТТПК и оптимизация принципов ведения сводного тематико-типологического плана комплектования.

Работа автоматизированной системы ведения сводного ТТПК была организована следующим образом.

Главное окно программы отображает информацию о двух базах данных. База данных рубрик УДК содержит записи вида «Индекс УДК – рубрика УДК (формулировка)». База данных библиотек содержит записи вида «Индекс УДК – тип, важность».

Связи этих двух баз данных организованы таким образом, что для каждого индекса (для каждой рубрики) выводится информация о том, какие библиотеки получают литературу по определенной теме и какие темы (в виде индексов) вообще комплектуется та или иная библиотека.

Источниками информации для базы данных рубрик УДК служили: рабочие таблицы классификации БЕН; формулировки тем библиотек ЦБС, к которым специалисты-систематизаторы всякий раз индивидуально подбирали индексы; составные индексы, присылаемые библиотеками, к которым специалисты-систематизаторы подбирали формулировки.

Источниками информации для базы данных библиотек были индексы тем, присылаемые библиоте-

ками, а также индексы, проставленные специалистами-систематизаторами на формулировки библиотек.

Важно отметить, что взаимодействие библиотек и специалистов-систематизаторов опосредовалось отделом комплектования отечественной литературы БЕН РАН. Все уточнения систематизаторами делались через сотрудников отдела комплектования. Другими словами, между систематизаторами и библиотеками ЦБС не было *обратной* связи.

Это привело к таким последствиям:

а) одна и та же тема могла скрываться под разными индексами. Причина этого – систематизатор каждый раз «с нуля» работал со списком тем каждой библиотеки, а индексирование в УДК – многоаспектно;

б) под одним индексом могли собираться разные темы. Причина этого – самостоятельное, но неверное индексирование библиотеками своих тем;

в) потеря части информации о темах. Причин этого две. Первая – невозможность для систематизатора передать индексом УДК формулировку темы. Вторая – невозможность для систематизатора дать формулировку для составного индекса, присланного библиотекой. В итоге такие индексы чаще всего игнорировались, и информация об этих темах просто не попадала в БД;

г) большинство индексов «пустые», невостребованные – они не несут информации о комплектуемой библиотеке.

Как показала работа сложившейся системы, проблема не может быть решена за счёт организации прямой и обратной связи между библиотеками сети и систематизаторами, поскольку *план комплектования* – рабочий инструмент *комплектователя*. Цели индексирования тем для плана комплектования иные, чем индексирование литературы для систематического каталога. В частности, в работе комплектователи ориентируются по индексам УДК, проставленным на книгах. Естественно, эти индексы не связаны с традицией систематизации в БЕН.

При анализе возможных путей выхода из сложившейся ситуации мы увидели, что при простом редактировании записей двух баз данных новые ошибки накладываются на старые ошибки. Таким образом, поддержание сложившейся системы требует огромных затрат времени и необходимости одновременной работы нескольких специалистов: комплектователя, систематизатора, работника сетевой библиотеки.

Поэтому нами были предложены новые принципы ведения сводного ТТПК БЕН. Ключевым моментом новых принципов является то, что предложена единая лингвистическая основа для всех участников процесса комплектования и ведения ТТПК (комплектователей, систематизаторов, библиотек ЦБС). Её использование позволяет изменить оба выявленных нами проблемных участка: источники информации и взаимодействие подразделений Централизованной библиотечной системы.

В соответствии с новыми принципами, база данных рубрик УДК имеет единственный источник – нормативную (официальную) УДК. В частности, это

позволяет комплектатору работать с индексами, проставленными на книгах.

Глубина этого рубрикатора индивидуальна для разных разделов УДК. Общий принцип здесь – отражение научных направлений (разделов конкретных наук) и ключевых понятий конкретной науки. Частные понятия и явления отсекаются. Это позволяет уменьшить объём базы данных, не потеряв при этом точности определения тем книжных изданий.

Темы комплектования библиотек в новой системе – это темы научных направлений, которые разрабатывает конкретный институт. Однако первые шаги в реализации новых принципов показали, что официальных формулировок, должным образом утверждённых и содержащихся в официальных документах, оказывается недостаточно в силу того, что иерархия понятий в Универсальной десятичной классификации иногда отличается от общепринятой. Поэтому предложено, чтобы формулировки тем сопровождалась списком ключевых понятий – ключевых слов – или аннотацией. Делается это для того, чтобы систематизаторы могли отразить индексами разделов наук эти темы, не потеряв при этом информацию.

Перед систематизаторами ставится теперь принципиально иная задача. Если раньше она звучала как «заиндексировать тему», то теперь – «указать индексы научных направлений, которые имеют отношение к конкретной формулировке научно-исследовательской темы и научно-исследовательской работы».

В качестве примера приведём направление «Исследования в области создания энергоресурсосберегающих и энергоэффективных технологий». Прямого индекса в УДК нет. Адаптируя формулировку этого направления для ТТПК, следует помнить, что в данном случае целью индексирования является не однозначное определение места этой темы в структуре знания, а *подбор источников, которые могут быть полезны при работе над данной темой*.

Поэтому в ТТПК обсуждаемое направление будет отражено следующими индексами:

66.012.37 Расход энергии

662.6/.8 Топливо в технике, в целом

502.173/.174 Управление природными ресурсами, рациональное использование, экономия ресурсов.

При таком подходе к формированию БД рубрик УДК мы не отказываемся от рубрик, за которыми не стоит ни одной библиотеки, хотя среди этих рубрик и остаются только рубрики высокого уровня иерархии. Такие «пустые» рубрики играют справочную роль для комплектатора, выступают в роли ключевых слов, расшифровывающих содержание того или иного научного направления для сотрудника, не являю-

щегося специалистом в конкретной науке. Это важно в тех случаях, когда на издании проставлен явно ошибочный индекс УДК, и определение темы книги приходится вести по названию, оглавлению, издательской аннотации и т.п.

Таким образом, новые принципы ведения сводного тематико-типологического плана комплектования БЕН РАН заключаются в следующем:

единые принципы индексирования тем;  
единые принципы определения формулировок тем;  
использование УДК как справочного аппарата для комплектатора по разным отраслям науки.

## СПИСОК ЛИТЕРАТУРЫ

1. ГОСТ 7.90–2007 Универсальная десятичная классификация. Структура, правила ведения и индексирования. – М.: Стандартинформ, 2008. – 23 с.
2. Дмитриева З.Г., Докторов Я.Я. Использование базы данных ТТПК в технологии комплектования БЕН РАН // Современные технологии в информационно-библиотечном обеспечении научных исследований: сб. научных трудов / под ред. Н.Е. Каленова. – М.: БЕН РАН, 1999. – С. 116–118.
3. Дмитриева З.Г., Докторов Я.Я. Развитие системы формирования тематико-типологических планов комплектования (ТТПК) // Новые технологии в информационном обеспечении науки: сб. научных трудов / под ред. Н.Е. Каленова. – М.: Биоинформсервис, 2001. – С. 118–119.
4. Кочукова Е.В., Наумова В.С., Павлова О.В. Традиции и современность в комплектовании библиотек ЦБС БЕН РАН отечественной литературой // Библиотека по естественным наукам РАН. Итоги и перспективы: сборник статей. – М., 2008. – С. 98–111.
5. Бочарова Е.Н., Кочукова Е.В., Докторов Я.Я. Актуализация сводного тематико-типологического плана комплектования ЦБС БЕН РАН // Библиосфера. – 2009. – №2. – С. 87–89.
6. Сводный тематико-типологический профиль комплектования единого библиотечного фонда БЕН АН СССР. – М.: БЕН АН СССР. –1990. – 302 с.
7. Бочарова Е.Н., Ивановский А.А. Формирование базы данных УДК для автоматизированной системы ведения сводного тематико-типологического плана комплектования Централизованной библиотечной системы БЕН РАН // Библиосфера. – 2013. – № 2.

Э. Р. Сукиасян (Российская государственная библиотека)

## Проблемы классификационной практики. Ситуационный анализ организационных и технологических решений

*Освещены проблемы классификационной практики в универсальных научных библиотеках*

**Ключевые слова:** *Средние таблицы ББК, универсальные научные библиотеки, технологические решения, проблемы классификационной практики, электронный каталог, сводный каталог*

Современная классификационная практика характеризуется немислимой пестротой принятых непродуманных решений. Более или менее удалось упорядочить положение дел в центральных библиотеках субъектов Федерации, ранее находившихся в ведении Министерства культуры (теперь эти библиотеки подчинены органам культуры области, края, республики). Поскольку речь идёт об универсальных научных библиотеках, то мы будем называть их сокращённо УНБ. В УНБ России должны использоваться Средние таблицы ББК. Другие решения неправомерны. У меня нет уверенности, что все УНБ начали перевод каталогов на Средние таблицы ББК. По разным причинам, чаще всего – демагогическим, некоторые «ждут». Администрация оказывается неспособной не только принять, но и провести в жизнь решения. Чаще всего при обсуждении принимается во внимание тот факт, что Средние таблицы ББК полностью не опубликованы («Вот когда закончится издание, мы начнём работу»). Несколько библиотек с миллионными фондами, каталоги которых были построены по полным таблицам ББК, долго в упор не видели нового издания (выходящего, между прочим, с 2001 г.), ожидая переиздания полных таблиц. Напомню: решение о том, что они издаваться не будут, было принято более десяти лет назад. Не изучив Средние таблицы, не проведя оценки их потенциальных возможностей, было заявлено: их детализации «нам никогда не хватит». Часть откровенных демагогов прямо заявляла: «эти таблицы для больших массовых библиотек, а не для УНБ».

Я мог бы вовсе не говорить «об оппозиции», но своих противников надо знать в лицо. Мы в Научно-исследовательском центре развития ББК (НИЦ ББК) их знаем. Давайте подумаем: из таких огромных по фондам книгохранилищам, получающим обязательный экземпляр, как, например, ГПНТБ СО РАН в Новосибирске или Дальневосточная государственная научная библиотека в Хабаровске, мы протестующих писем не получаем. Идёт нормальная работа. Мы сразу же заявили: если библиотеки увидят, что Средние таблицы недостаточны для фондов библиотеки, НИЦ ББК обеспечит каждую такую библиотеку – по

соответствующим разделам – полными таблицами. Эталон полных таблиц ведётся в НИЦ ББК на основе развития Средних таблиц. Только покажите нам, где именно недостаточен объем Средних таблиц.

Хочу подчеркнуть: в Средних таблицах глубина детализации примерно в два раза перекрывает таблицы Десятичной классификации М. Дьюи (4 тома). В ББК гораздо выше комбинационные возможности: Средние таблицы обеспечены системой типовых делений общего применения (ТДОП) и специальных типовых делений. По нашему мнению, в детализации могут нуждаться библиотеки узкоспециальных НИИ, но не УНБ.

Многие библиотеки до сих пор не внедрили в систематические каталоги принципы обратнотронологической расстановки карточек, продолжают строить каталоги в алфавитном порядке записей внутри каждого деления [1-2]. Удивительная память – помнят до сих пор отжившую норму: за разделителем должно быть 40-70 карточек, не более. Такой нормы сейчас нет. Нормой стала простановка на карточке и в записях Электронного каталога (ЭК) даты завершения обработки. А для ЭК должно быть нормой представление информации на экране в обратнотронологическом порядке (именно так: по этой дате), а не «вываливание» массива на пользователя в произвольном порядке.

Мы не возражаем, чтобы библиотеки, если они найдут силы и время, отредактировали бы свои каталоги. Но мы считаем, что во многих случаях редактирование невозможно: полностью изменилась структура ряда разделов. Не перестаю повторять: у нас в 1991 г. изменился государственный строй, стала другой идеология общества. У нас стали иные экономика и политика. Да, возможно подразделения искусства, физкультуры и спорта, архитектуры, литературоведения можно отредактировать, строя единый каталог для литературы за много лет. Но нельзя собирать литературу по марксистско-ленинской экономике или педагогике, эстетике и в том же каталоге – современную литературу. Нет сегодня «Научного коммунизма», и не должно быть такой «науки» в современном каталоге. Мы пошли

на коренные изменения в ряде отделов и разделов, имея в виду: единственное решение для библиотеки – открыть новый ряд каталога.

Мне пишут: нашим читателям очень трудно. У нас и так три ряда сводного каталога (СК), вы предлагаете организовать четвёртый. Отвечаю: сначала надо разобраться, почему у вас не выполнены приказы Министерства 1960-х гг. о переводе на ББК... А вот зарубежный опыт свидетельствует: в некоторых странах каждые 10 лет организуется новый ряд СК со своим указателем. Там считают, что научная терминология и общественное развитие за 10 лет делают шаг вперёд, изменяется терминология. В Германии мне так и сказали: сам «орднунг» в каталоге должен быть другой, мы и систему кардинально меняем... Вот там, где нет демагогии, есть «орднунг», т.е. порядок. А у нас кое-где порядка нет.

Давайте договоримся: там, где я буду говорить о ББК, я буду иметь Средние таблицы. Если вы в своей библиотеке ещё не освоили методику и технологию перехода, работайте – другого пути нет. Ждать нечего. Мне понятно: многие тянут время до своего ухода на пенсию. Так ушли бы! Нет, наступает время, но они не уходят, продолжая играть в свою дудку. И их руководители, судя по всему, не могут принять необходимое в таких случаях решение.

Как мне хотелось бы сказать, что с УДК дело обстоит гораздо лучше! Что все перешли на последнее русское издание, отредактировали каталоги, своевременно, раз в год вносят дополнения и исправления, ведут указатели... Если бы была хоть какая-нибудь уверенность. У нас ведь нет сейчас по сути дела ни методической помощи на местах, ни, тем более, какого-либо управления и контроля. Во время своих поездок я захожу не только в те библиотеки, которые работают по ББК. Во многих вузах, например, часть каталогов организована по УДК, другая часть – по ББК. Хочется открыть и посмотреть, просто увидеть, как реализуется огромная, поистине героическая работа небольшого коллектива ВИНТИ РАН по переводу и изданию таблиц УДК. Я ведь продолжаю любить УДК, непосредственно занимаясь ББК. З.Н. Амбарцумян говорил: специалист по классификациям должен знать по возможности все, которые применяются. Хотя бы для того, чтобы объективно оценивать достоинства и недостатки каждой из них. Стараюсь. Поэтому сразу вижу в каталогах то, что тщательно от меня скрывают. Глаз выхватывает индексы с определителем точки зрения, другие фокусы, от которых мы освободились... Почему не редактируете? Иногда ответы совпадают с теми, о которых я рассказывал выше.

Поэтому хочу сказать «давайте договоримся» второй раз: если уж вы взяли на себя ответственность за каталог по УДК, то ведите его так, чтобы не было стыдно показать специалистам, которые знают современные таблицы УДК.

После этого болезненного предисловия перехожу к ситуациям, которые случаются на практике. Хочу сразу предупредить: все они реальны, взяты из писем, которые приходят ко мне. Некоторые поступают к главному редактору ББК, другие пишут члену ред-

коллегии журнала «Научные и технические каталоги», обращаются просто к докладчику на наших семинарах и конференциях, преподавателю со стажем. Со многими библиотекарями иногда долго продолжается переписка. Приходится убеждать.

Традиционная ситуация: в библиотеке параллельно и одновременно используются как УДК, так и ББК. Сразу же хочется сказать: юристы подсказали мне, что ссылки на те или иные Постановления Совмина (например, известное № 445 от 11.05.1962) уже неправомерны. Как и ведомственные решения, принятые до 1991 г. У нас действует лишь ГОСТ 7.59–2003 в п. 5.2.7 которого сказано «В зависимости от вида документа и поставленных задач следует применять одну или несколько универсальных систем классификации - классификационных ИПЯ». В списке классификаций, подчеркну – на равных и без указаний «где какие», перечислены все, в частности и ББК, и УДК. В конце сказано: «При необходимости допускается применять также и локальные (отраслевые, специальные) классификационные ИПЯ».

Хочу попросить: если вам известны какие-либо ведомственные нормативные акты, регламентирующие применение классификационных систем, принятые после 1991 г., сообщите мне о них. Буду разбираться в правомочности органа, который хочет поставить себя выше Государственного стандарта.

Решение о границах применения УДК и ББК в библиотеке принимается её директором. Оно предварительно обсуждается на Совете по системе каталогов. Я не возьму на себя ответственность определять какие-либо границы. И не только потому, что являюсь – по должности – лицом заинтересованным. Библиотека должна опираться на мнение квалифицированных экспертов (например, в вузе – профессоров и преподавателей соответствующих кафедр). Дайте им проанализировать путём сравнения таблицы, попросите ответить на чётко поставленный вопрос.

Ещё примерно 10-15 лет мы не могли выбрать систему, наиболее полно отражающую современную литературу. Для вузов таким критерием будет соответствие «языка» классификационной системы языку учебных программ вуза. Сравняйте, выбирайте. Конечно, иметь в Сводном каталоге одну систему желательно, технологически удобнее. Средние таблицы ББК по технике будут изданы летом 2013 г., через год будут опубликованы таблицы по естественным наукам. УНБ пока работают по Таблицам для областных библиотек. А ещё через год библиотеки получат совершенно новый 1-й выпуск Средних таблиц – отдел «Междисциплинарное и общенаучное знание» [3].

Если библиотека работает одновременно с УДК и ББК, серьёзных технологических проблем не возникает. Главная задача – чётко провести границы. Индексационные различия позволяют: в УДК, как известно, отделяются точкой каждые три знака, в ББК – первые два, далее каждые три. От русских букв в основных делениях ББК надо было давно освободиться. Я бы рекомендовал строить единый алфавитно-предметный указатель к СК (АПУ). В библиотеке должен сохраняться карточный СК, его части, по-

строенные по УДК и ББК, надо выделить во внешнем оформлении, например цветом этикеток на ящиках. Придётся поработать и со средствами наглядной информации. Читателю должно быть понятно, где какой индекс искать (если он начал с АПУ), где в каталоге, какая отрасль знания (если он работает с внешним оформлением).

О том, почему надо сохранить карточный СК при наличии ЭК, я писал неоднократно [3]\*. Никому до сих пор не удалось доказать, что ЭК полностью может заменить карточный СК, исследований не проводилось. Повторяться не буду (тема другая), обратитесь к литературе. Добейтесь с программистами: в ЭК должен быть поиск по индексам сверху вниз (т.е. иерархический), а также поиск по элементам классификационного индекса (например, определителям или типовым делениям). Таблицы классификации должны быть доступны читателям в виде самостоятельной базы знаний. В октябре 2006 г. с таким предложением ко всем разработчикам автоматизированных библиотечных информационных систем (АБИС) обратились участники Всесоюзного совещания в Москве. Это обращение опубликовано в печати и направлено всем разработчикам АБИС отдельным письмом [4].

Усложним ситуацию: допустим, что библиотека редактирует СК или строит его новый ряд. Алгоритм разработан. Обратитесь к нашему пособию «Новые таблицы ББК. Организация и технология использования» [5]. Оно адекватно «работает» и в ситуациях, связанных с УДК: технология та же. Пособие можно использовать во всех случаях, когда перед библиотекой встает задача перехода от двух систем к одной, единой.

Рассмотрим ситуации, которых раньше у нас не было: объединяются библиотеки с имеющимся в каждой своим справочным аппаратом. Сегодня в стране идёт массовое объединение вузовских библиотек, создаются сильные университеты и поэтому ставится задача слияния нескольких, ранее самостоятельных библиотек в единую библиотеку – систему.

Мне кажется, что в качестве первого шага надо освободить фонды от устаревшей и мало (редко) спрашиваемой литературы, уменьшить экзemplарность тех изданий, которые в учебном процессе не используются. Поймите меня правильно: если раньше мы вынуждены были держать фонды, чтобы сохранять количественные показатели для относительных измерений (работал показатель единиц фонда на одного студента), то теперь этого делать не надо: критерии изменились. Докажите руководителям вуза, что книгохранилища перегружены литературой, которая в учебном процессе не будет никогда использоваться. Один-два экземпляра сохраняются для научных целей. Добейтесь сначала полной ликвидации позорного для отечественных библиотек явления, которое мы понимаем, а за рубежом понять не могут: «штабелирование фондов».

Конечно, все моменты движения фондов должны сразу же отражаться в системе каталогов, начиная с

генерального (служебного) алфавитного каталога. Не ошибусь, наверное, если скажу: никто в это не верит. В спешке ведь всё делается. О каталогах можем и не вспомнить вовсе, тем более, что фондами занимаются у нас одни сотрудники, а каталогами другие. Библиотеку как систему, в которой все элементы взаимосвязаны, пока мы не воспринимаем. Поэтому вторым действием должна быть сверка фондов с каталогами. Технология проведения такой сверки разработана, алгоритм описан мною [2, с. 157-160].

Справились? Теперь пришла пора оценить ситуацию с классификационной стороны. В какой библиотеке как построен СК, по какой системе классификации (по какому изданию), насколько состав и содержание СК отражает карточный алфавитно-предметный указатель (АПУ). Последнее станет ясно, если вы проанализируете технологию его ведения. Понятно, что говорить о «слиянии» можно лишь в том случае, если эта операция свершается с качественно равноценными элементами. Иначе говоря, после принятия решения работа ведётся отдельно в каждой библиотеке. На этом этапе может очень помочь ведение единого на все фонды ЭК. Никакой конверсии не требуется! Один из ЭК признаётся центральным. Записи из других каталогов сверяются. Если такая запись есть – дополняются холдинговые (фондовые) данные в формате. Если записи нет – она оформляется (переносится) в центральный каталог.

Если объединяются библиотеки, находящиеся в одном городе, обязательно организуется центральный справочно-библиографический отдел, в котором и находятся центральные каталоги. При этом в каждой из библиотек (в зданиях нескольких или отдельных факультетов, на кафедрах) должен быть организован удалённый доступ к центральному ЭК. На фонд «своей» библиотеки (как части системы) может организовываться карточный СК с АПУ.

Сложнее решать аналогичные проблемы, если государство приняло решение об объединении вузов, территориально находящихся в разных городах. Конечно, единого справочно-библиографического обслуживания в одном месте уже не получится: каждый студент имеет право на получение библиотечных и библиографических услуг в полном объёме по месту учёбы и проживания. Единый центральный ЭК можно предоставить. Проблемы нет. Лучше выставить его в Интернете. Ещё лучше – поставить более полезную, но несоизмеримо сложную задачу: выставить в Интернете единый ЭК библиотек города, области. Не научились мы этому, не умеем. Даже не видим необходимости. Её постоянно ощущают наши читатели, жители регионов. Но мы их не спрашиваем, они у нас приучены не задавать лишних вопросов. Даже подумать страшно, какую обструкцию устроили бы нам американские студенты, если бы мы их лишили столь привычных услуг: открытых фондов к книгохранилищам в миллионы томов, свободного и бесплатного доступа к информации в журналах и сборниках, работающих круглосуточно библиотек, Wi-Fi не только в университетах и общежитиях, но и в столовых, ночных клубах и автобусах...

\* Указанной проблеме посвящен ряд материалов [2, с. 219-232].

Нам придётся много работать. Строить планы и графики, научиться считать время и деньги, оценивать качество, функциональную и экономическую эффективность каждого шага. Каталоги – это только начало.

## СПИСОК ЛИТЕРАТУРЫ

1. Организация работы республиканской (АССР), краевой, областной универсальной научной библиотеки. Ч. 1. Формирование библиотечного фонда. СБА. Технология каталогизации : метод. рекомендации. – М., 1991. – 127 с.
2. Сукиасян Э.Р. Каталогизация. Классификация. Электронные каталоги и автоматизированные библиотечные системы : избр. статьи. – СПб : Профессия, 2009. – 535 с.
3. Междисциплинарное и общенаучное знание: классификационный анализ. Подходы к формированию структуры и содержания нового отдела ББК // Науч. и техн. б-ки. – 2010. – № 3. – С. 69–78; № 4. – С. 49–57.
4. О Всероссийском совещании библиотек по вопросам ББК (Москва, 10–12 октября 2006 г.) // Университетская книга. – 2006. – № 12. – С. 50–53; Библиотека. – 2006. – № 11. – С. 63–65; Библиотекведение. – 2007. – № 2. – С. 120–125; Информ. бюл. РБА. – 2007. – № 44. – С. 92–100; Науч. и техн. б-ки. – 2007. – № 4. – С. 100–114.
5. Сукиасян Э.Р. Новые таблицы Библиотечно-библиографической классификации : организация и технология использования : метод. рекомендации. – М. : Либерея, 2003, допечатка тиража 2005. – 95 с.

УДК 025.45.05:[002:004]

М.К. Сухоруков, А.А.Гребенникова (Научно-технический центр «Информрегистр»)

## Тематическая рубрикация электронных изданий средствами УДК: практика применения, проблемы и перспективы

*Описана система формирования Депозитария электронных изданий, которую ведет Научно-технический центр «Информрегистр» (НТЦ «Информрегистр»).*

**Ключевые слова:** электронное издание, обязательный экземпляр электронного издания, депозитарий электронных изданий, рубрикация электронных изданий

Российское законодательство предписывает обязательный библиографический учет для электронных документов на съемных носителях, которые отнесены к категории обязательного экземпляра и должны входить в состав национального библиотечно-информационного фонда. Национальный библиотечно-информационный фонд документов Российской Федерации – это собрание всех видов документов, комплектуемое на основе обязательного экземпляра, распределяемое в соответствии с Федеральным законом Российской Федерации «Об обязательном экземпляре документов» от 29 декабря 1994 г. № 77-ФЗ\*, предназначенное для постоянного хранения и общественного использования и являющееся частью культурного достояния народов Российской Федерации. НТЦ «Информрегистр», в соответствии с Федеральным законом № 77-ФЗ, участвует в исполнении установленных законодательством функций по госу-

дарственному библиографическому учету обязательного экземпляра электронных изданий. Массив электронных изданий формируется Депозитарием электронных изданий – структурным подразделением НТЦ «Информрегистр». В настоящее время Депозитарий содержит свыше 28 тыс. наименований обязательных федеральных экземпляров электронных изданий (рис. 1). Сведения о них представлены в электронном каталоге «Российские электронные издания» на сайте [www.inforeg.ru](http://www.inforeg.ru) в свободном доступе.

Государственный библиографический учет подразумевает постоянную работу по библиографическому описанию и классификации электронной продукции, поступающей в Депозитарий. В настоящее время рубрикация обязательных федеральных экземпляров электронных изданий в НТЦ «Информрегистр» производится на основе Универсальной десятичной классификации (УДК) и Государственного рубрикатора научно-технической информации (ГРНТИ).

\* Федеральный закон Российской Федерации от 29 декабря 1994 года № 77-ФЗ "Об обязательном экземпляре документов" (в ред. Федеральных законов от 11.02.2002 № 19-ФЗ, от 27.12.2000 № 150-ФЗ, от 24.12.2002 № 176-ФЗ, от 26.03.2008 № 28-ФЗ).

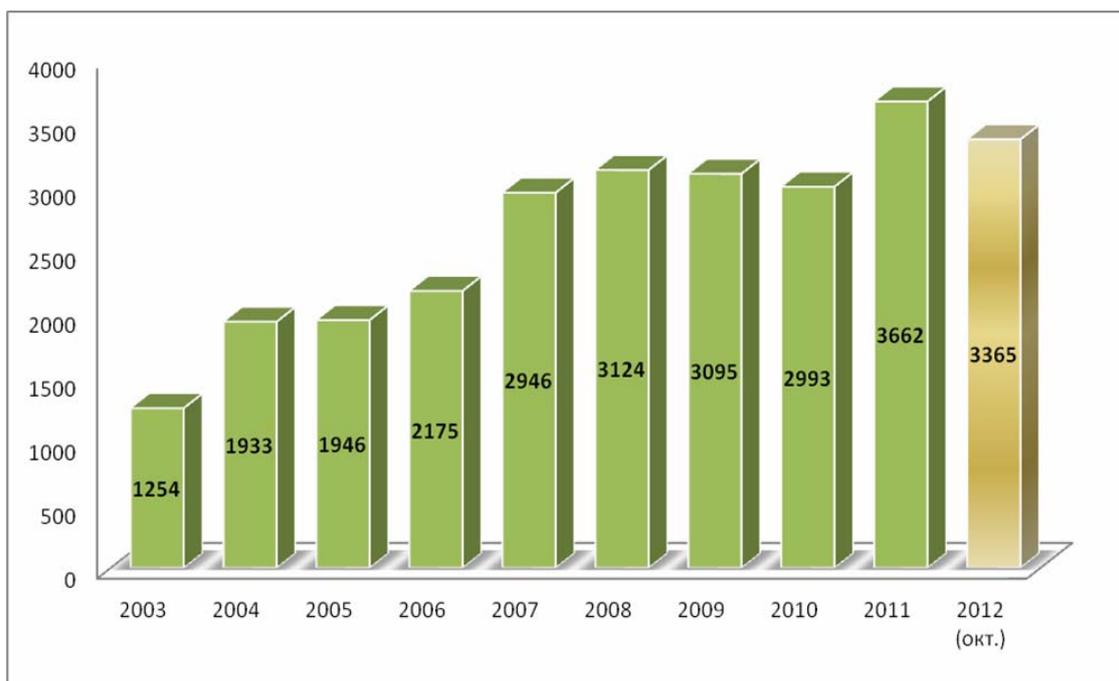


Рис. 1. Динамика поступлений электронных изданий в каталог Депозитария электронных изданий НТЦ «Информрегистр» (2003г.– октябрь 2012г.)

Исторически, до внедрения автоматизации процессов регистрации обязательных федеральных экземпляров электронных изданий, в НТЦ применялась рубрикация на основе УДК. Глубина рубрикации не превышала второго-третьего уровня иерархии классификатора; система определителей не использовалась. Это было связано с превалированием всё же регистрирующей, а не библиографической функции НТЦ «Информрегистр».

В 2003 г. в организации была внедрена автоматизированная библиотечная информационная система «ИРБИС». АРМ «Каталогизатор» пришлось существенно доработать с учетом специфики регистрации и библиографического описания электронных ресурсов. Рубрикация поступающих электронных изданий начала осуществляться также с помощью Государственного рубрикатора научно-технической информации (ГРНТИ). В качестве методического обеспечения использовалась «Таблица перекодировки УДК – ГРНТИ». Рубрикация средствами УДК стала полнее за счет использования более дробных индексов, а также общих определителей формы документов и общих определителей места. Появилась также возможность проводить координатное индексирование с помощью ключевых слов.

С декабря 2010 г. государственная регистрация и учет электронных изданий осуществляются в Автоматизированной информационной системе государственной регистрации информационных ресурсов (СГРИР), получившей свидетельство о соответствии выгружаемых библиографических записей формату RUSMARC. Переход на СГРИР был обусловлен недостаточным соответствием возможностей АБИС «ИРБИС» задачам и целям НТЦ «Информрегистр»

по формированию Депозитария электронных изданий. Основными функциями СГРИР являются: ввод данных об электронных изданиях и организациях-производителях, редактирование данных, ведение внутрисистемных классификаторов и справочников, вывод информации (в режиме поисковых запросов и стандартных форм вывода), в том числе вывод статистических данных, поиск по метаописаниям электронных изданий и их производителей, контроль за выполнением процесса регистрации.

Автоматизированная информационная система государственной регистрации информационных ресурсов позволила встроить «дерево» рубрик ГРНТИ, что существенно ускорило процесс рубрикации электронных изданий. Для введения классификационных индексов в СГРИР предусмотрены поля, соответствующие полям RUSMARC для индексов УДК и ГРНТИ. Также в НТЦ «Информрегистр» используется «База данных УДК. Сокращенные рабочие таблицы» (разработчик – ГПНТБ России). В практике индексирования стали в обязательном порядке использоваться общие определители формы документов; чаще стали применяться общие определители места. Так, обязательно используются такие определители формы документов, как (075) – учебная литература, (035) – справочная литература, (035.3) – монографии, (063) – издания конференций и т.д.

В НТЦ «Информрегистр» поступают обязательные федеральные экземпляры электронных изданий по всем отраслям знания, поэтому фонд Депозитария можно назвать универсальным. Распределение обязательных федеральных экземпляров электронных изданий по рубрикам УДК по состоянию на ноябрь 2012 г. представлено на рис. 2.

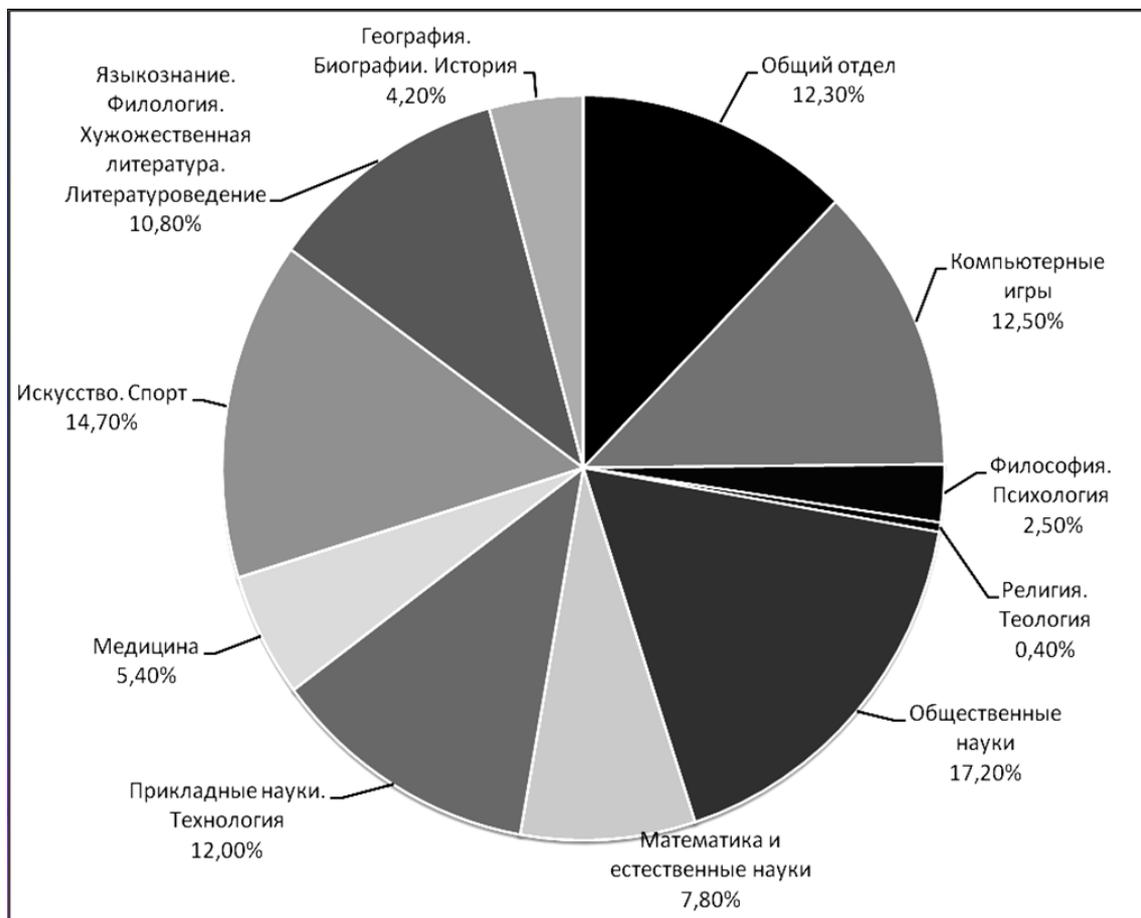


Рис. 2. Распределение обязательных федеральных экземпляров электронных изданий по рубрикам УДК в фонде Депозитария электронных изданий НТЦ «Информрегистр» (1996 г. – ноябрь 2012 г.)

Специфика носителя информации – машиночитаемые диски CD/DVD – определила поступление такого потока изданий, как компьютерные игры. Наибольшее количество регистрируемых компьютерных игр пришлось на начало 2000-х гг. В связи с этим возникла необходимость выделить и более дробно систематизировать этот поток электронных изданий, а значит доработать соответствующие разделы УДК. Для решения этой проблемы был использован принцип УДК, предоставляющий право систематизатору самостоятельно вводить алфавитные (именные, буквенные) подразделения. До 2003 г. электронные издания, содержащие компьютерные игры, относили к разделу УДК 790.034 «Компьютерные игры», с последующим дробным делением по жанрам игр. В 2003 г., в период начала автоматизации процессов регистрации и депонирования обязательных федеральных экземпляров электронных изданий, был доработан раздел УДК 004.9 «Прикладные информационные (компьютерные) технологии». На рис. 3 приводится список разделов УДК, которые использовались в НТЦ «Информрегистр» для рубрикации потока электронных изданий, содержащих компьютерные игры.

В настоящее время задачи по рубрикации обязательных федеральных экземпляров электронных изданий средствами УДК, стоящие перед научно-техническим центром, расширяются. Это обусловлено необходимостью исполнения недавнего соглашения

между ФГУП НТЦ «Информрегистр» и ФГБУН «Российская книжная палата». С целью расширения информирования неограниченного круга пользователей о зарегистрированных обязательных экземплярах электронных изданий, во исполнение ст. 22 Федерального закона «Об обязательном экземпляре документов», в еженедельном библиографическом указателе «Книжная летопись» начали ежемесячно размещаться сведения об электронных изданиях, зарегистрированных во НТЦ «Информрегистр». В частности, для организации этого процесса в Автоматизированной информационной системе государственной регистрации информационных ресурсов (СГРИР) было добавлено поле «Индекс УДК РКП». От ФГБУН «Российская книжная палата» получен список рубрик УДК, которые используются для формирования старейшего государственного библиографического указателя «Книжная летопись». Первая партия библиографических описаний обязательных федеральных экземпляров электронных изданий опубликована в № 44 «Книжной летописи» за текущий год.

На основе договоров с национальными библиотеками России НТЦ «Информрегистр» передает им библиографические записи в формате RUSMARC. Специализированные отделы систематизации и предметизации национальных библиотек в случае необходимости используют их как базовые для углубленной систематизации.

790.034 Компьютерные игры (до 2003 г.)	004.9 Прикладные информационные (компьютерные) технологии (после 2003 г.)
790.034.1 Энциклопедии игр	004.9 аркады
790.034.10 Квесты	004.9 дополнения
790.034.11 Экшены	004.9 искусство
790.034.12 Симуляторы	004.9 квесты
790.034.13 Стратегии	004.9 логические игры
790.034.14 Приключения	004.9 менеджеры
790.034.17 Эротические	004.9 приключения
790.034.18 Ролевые игры	004.9 развивающие игры
790.034.19 Фэнтэзи	004.9 ролевые игры
790.034.2 Сборники игр	004.9 сборники игр
790.034.3 Универсальные игры	004.9 симуляторы
790.034.4 Дополнения. Утилиты. Моды, карты, патчи и пр.	004.9 стратегии
790.034.5 Обучающие игры	004.9 универсальные игры
790.034.6 Азартные игры	004.9 экшн
790.034.7 Логические игры	004.9 энциклопедии игр
790.034.8 Спортивные игры	004.9 эротические
790.034.9 Аркады	

Рис. 3. Детализация разделов УДК для потока электронных изданий, содержащих компьютерные игры

Таким образом, в практике рубрикации электронных изданий средствами УДК используются основные методические принципы применения этой классификационной системы. Не стремясь к дробному индексированию, используя вторые-третьи уровни иерархии классификатора и две таблицы определителей (формы документов и

определителей места), ориентируясь на главную – регистрирующую – функцию в своей деятельности, НТЦ «Информрегистр», тем не менее, может представить объективную картину распределения обязательных федеральных экземпляров электронных изданий по отраслям знания.

УДК 025.4

В. Н. Белозеров, А. В. Шапкин (ВИНИТИ РАН),  
Н. Н. Шабурова (Институт физики полупроводников им. А.В. Ржанова СО РАН)

## Табличное и тезаурусное сопоставление классификационных систем

*Рассматриваются особенности и содержание двух систем сопоставления разнородных классификационных систем: Системы классификационных схем ВИНИТИ РАН и Тезауруса тематических рубрик Института физики полупроводников СО РАН.*

**Ключевые слова:** тезаурус тематических рубрик, классификационные схемы, классификаторы, дескрипторы

В большинстве информационных и библиотечных систем приходится использовать систематизацию информации не по одной классификационной схеме, а одновременно по нескольким для выполнения различных технологических задач. Информационный

фонд может комплектоваться из внешних источников, упорядоченных по своим классификациям, не совпадающим с систематизацией фонда, а пользователи фонда могут требовать поставки информации, отобранной и упорядоченной по другой схеме, опять

же не совпадающей с систематизацией фонда. При этом возникает потребность навигации по классификационным схемам и перевод рубрик одного классификатора в рубрики другого классификатора. Наличие соответствующего аппарата может быть полезным для организации мультибазового поиска документов и для реализации многоаспектного доступа к информации в рамках одного ресурса.

В ВИНТИ РАН реализована табличная система сопоставления классификаций, состоящая из набора таблиц классификационных систем, в которых для каждой табличной записи указаны близкие по смыслу рубрики одной или нескольких других классификаций [1]. В Систему классификационных схем (СКС) сведено несколько важных для нас классификаций иерархического типа, используемых в крупных центрах научно-технической информации, в сфере образования и культуры, экономической деятельности. Набор представленных классификаторов определяется, в первую очередь, потребностями обработки в ВИНТИ РАН потока научно-технической литературы по своему профилю, а также необходимостью осуществлять взаимодействие с национальными и международными информационными службами, научными и образовательными институтами, учреждениями культуры.

В настоящее время Система классификационных схем включает следующие классификаторы, указанные в таблице:

Для целей информационной навигации между классификаторами в СКС установлены связи — отображения рубрик по семантической близости представляемых понятий. Сопоставление отдельных классификационных схем выполняется экспертами Научно-методического отдела и Отделов научной информации ВИНТИ РАН, а также специалистами Отдела программных систем в результате статистической обработки баз данных. Предусмотрены следующие режимы пользования СКС:

- **Система классификационных схем** в режиме общего доступа открывает ограниченный набор классификаторов из представленных в СКС, а именно – ГРНТИ, Рубрикатора ВИНТИ, Номенклатуры ВАК, Классификатора РФФИ. Здесь доступно изучать связи рубрик данных классификаций с другими классификаторами.

- **Служба отображений рубрик** в ручном режиме предлагает услугу перевода кода рубрики одного классификатора в коды рубрик другого классификатора без просмотра самих классификаторов. Доступны рубрики всех классификаторов, представленных в СКС.

- **Служба отображений рубрик** в режиме программного доступа предлагает услуги навигации по классификациям в автоматическом режиме при обработке поискового запроса на виртуально объединенном пространстве разнородных информационных ресурсов.

Таблица

<i>Наименование</i>	<i>Дополнительные сведения</i>	<i>Отражено на:</i>
Journal of Economic Literature Classification –	Классификация Журнала экономической литературы (США)	ГРНТИ
Рубрикация журнала Library and Information Science Abstracts	Рефераты по библиотечно-информационной науке (Великобритания)	РВИНТИ
Mathematics Subject Classification, 2000 –	Математическая предметная классификация MSC (международная)	РВИНТИ
Государственный рубрикатор научно-технической информации	ГРНТИ (Россия)	РВИНТИ, УДК, ВАК
Классификатор Российского фонда фундаментальных исследований	РФФИ	ГРНТИ
Международная патентная классификация	МПК 7	РВИНТИ
Номенклатура специальностей научных работников (2002-2009 гг.) –	Высшая аттестационная комиссия России (ВАК)	ГРНТИ, РВИНТИ
Общероссийский классификатор видов экономической деятельности	ОКВЭД	ГРНТИ, РВИНТИ
Общероссийский классификатор стандартов	ОКС	РВИНТИ
Радиационная экология, радиогеоэкология, радиоактивное загрязнение (проекты ИГЕМ)	Радиогеоэкология – Россия	ГРНТИ, РВИНТИ
Рубрикатор ВИНТИ (текущий, 2013 г.) –	РВИНТИ, Россия	МПК, Радиогеоэкология, MSC, ОКС, ГРНТИ, ВАК
Универсальная десятичная классификация	УДК	ГРНТИ

По мере расширения сферы навигации и открытия доступа к новым ресурсам состав СКС будет пополняться новыми классификаторами, которые будут оснащаться взаимными отображениями рубрик. База данных Системы классификационных схем поддерживается в локальной сети Института на SQL-сервере Отделом программных систем. Возможна установка системы заинтересованным пользователям с учётом их потребностей в тех или иных классификационных системах.

В случаях, когда необходимо определить место новой тематики в различных классификациях, представляется удобным использовать предметный вход в сопоставительную систему. Такая возможность реализована в Институте физики полупроводников СО РАН (Новосибирск) совместно с ВИНТИ РАН. Система представляет собой тезаурус тематических рубрик нескольких классификационных систем по физике полупроводников [2], который используется для тематического индексирования и поиска публикаций. В настоящее время тезаурус пополняется тематикой нанотехнологий и электроники и будет охватывать всю проблематику исследований Института. В основе тезауруса лежат соответствующие разделы следующих классификаций:

- Универсальной десятичной классификации (УДК)
- Библиотечно-библиографической классификации (ББК)
- Государственного рубрикатора научно-технической информации (ГРНТИ)
- Рубрикатора отраслей знания ВИНТИ
- Схемы классификации по физике и астрономии (Physics and Astronomy Classification Scheme – PACS), США
- Классификаторов Федерального портала «Наноматериалы и нанотехнологии» (НАНО).

В качестве входных терминов служат полные тексты рубрик этих классификаций, а также выделенные из них ключевые слова. Большая часть терминов выступают как дескрипторы тезауруса, между которыми установлены логические связи на основе отношений включения и пересечения объёмов понятий. Указаны также в качестве синонимов альтернативные формулировки терминов. Всего в тезаурусе установлены перекрёстные соответствия около 1800 классификационных рубрик, а общее число элементов дескрипторных статей, фиксирующих связи терминов, превышает 9,5 тыс. Приведем примеры тезаурусных статей:

<b>радиационные эффекты в полупроводниках</b> ББК В379.212.7 PACS 61.82.Fk *УДК 537.311.322.04:538.97 С: <i>влияние излучения/облучения на структуру и свойства полупроводников</i> В: влияние облучения на свойства твёрдых тел В: структура полупроводников В: физические свойства полупроводников Н: взаимодействие ионов с полупроводниками Н: деканалирование заряженных частиц в полупроводниках Н: ионная имплантация полупроводников Н: каналирование и деканалирование заряженных частиц в полупроводниках Н: радиационный эффект в ферритах /полупроводниках/ А: излучения – действие на проводимость	<i>дескриптор</i>
	<i>индексы различных классификаций</i>
	<i>синоним</i>
	<i>вышестоящие (родовые) понятия</i>
	<i>видовые (частные) понятия</i>
	<i>близкое по смыслу понятие</i>

<b>физика наноструктур</b> ВИНТИ 291.19.22 ГРНТИ 29.19.22 *УДК 538.9-022.532 С: <i>наноразмерные объекты (физика)</i> В: физика твёрдых тел Н: свойства нанобъектов А: мезоскопические структуры А: низкоразмерные структуры
--

<b>атомнослоеое осаждение</b> ВИНТИ 291.19.22.13.11.15.11 НАНО09 2.1.1.3 *УДК 539.2-022.532::544.023.22-123 В: физические методы нанесения нанобъектов на подложку
--

## СПИСОК ЛИТЕРАТУРЫ

1. Доступ к Системе классификационных схем ВИНТИ. – URL: [http://www2.viniti.ru/index.php?option=com\\_content&task=view&id=134&Itemid=63](http://www2.viniti.ru/index.php?option=com_content&task=view&id=134&Itemid=63)

2. Доступ к тезаурусу ИФП. – URL: <http://lib.isp.nsc.ru/library/links.html?PHPSESSID=76hc9974fhtfumpas96fq34420>.

УДК 025.4.02:027.021

Н.Н. Шабурова (Институт физики полупроводников им. А.В. Ржанова СО РАН)

## Проблемы систематизации библиотечных фондов академического института

*На примере библиотеки Института физики полупроводников Сибирского отделения РАН описываются проблемы систематизации библиотечных фондов, создание и ведение электронного каталога в научной библиотеке.*

**Ключевые слова:** централизованная библиотечная сеть, комплектование, библиотечные фонды, электронный каталог, систематизация и каталогизация библиотечных фондов

Централизованная библиотечная сеть (ЦБС) СО РАН, которую возглавляет ГПНТБ СО РАН, состоит из центральных библиотек в столицах областных научных центров (Иркутск, Красноярск, Омск, Томск, Улан-Удэ, Якутск) и 65 академических библиотек при научно-исследовательских институтах (НИИ).

Библиотека Института физики полупроводников (ИФП) СО РАН — типичный представитель академических библиотек Сибирского региона, она создана и развивалась в одинаковых исторических, территориальных и экономических условиях, как и большинство библиотек других НИИ. ИФП ведет исследования и разработки как по естественнонаучному, так и по техническому направлениям. Основные из них следующие:

1) актуальные направления физики конденсированных сред, в том числе физика полупроводников и диэлектриков, физика низкоразмерных систем;

2) элементная база микроэлектроники, наноэлектроники, квантовых компьютеров, в том числе физико-химические основы технологий микроэлектроники, наноэлектроники, оптоэлектроники, акустоэлектроники, микросенсорики;

3) актуальные проблемы оптики, лазерной физики, включая квантовую электронику.

В соответствии с этими направлениями комплекуются и библиотечные фонды Института, которые на 01.01.2012 г. содержали 65 % естественнонаучных изданий и более 30 % технической литературы. Всего в библиотеке 77816 единиц хранения (в т.ч. 17226 на иностранных языках), в том числе по видам изданий:

25060 – книги (из них 1845 иностранные),

51479 – периодические издания (из них 15381 иностранные).

Эти фонды располагаются: в книгохранилище – 62973 экз., в читальном зале – 8928 экз., в «пассивном» (обменно-резервном) фонде – 5665 экз., в специальном

распределенном резервно-страховом фонде (создан на основании «Постановления Президиума СО РАН » от 21.06.2002 из трудов сотрудников НИИ) – 250 экз.

Следует отметить, что при создании ИФП в 1964 г. в результате слияния двух других НИИ («Физики твердого тела и полупроводниковой электроники» и «Радиофизики и электроники») библиотечный фонд последнего насчитывал свыше 10000 экземпляров, в том числе 6500 – книг, и был взят за основу для библиотеки нового института. Систематизировался фонд по Универсальной десятичной классификации (УДК). Так, с 1964 г. библиотека ИФП, как преемник, на протяжении 15 лет продолжала систематизировать и каталогизировать поступающие издания по этой же классификации.

В начале 80-х гг. прошлого столетия фонды и каталоги Библиотеки ИФП, как и многих других библиотек НИИ СО РАН, были конвертированы из УДК в Библиотечно-библиографическую классификацию (ББК). К тому времени книжные фонды насчитывали более 16000 единиц хранения, а алфавитный, систематический и предметный каталоги - около 100000 каталожных карточек. Эта работа велась параллельно с обработкой вновь поступающих изданий и заняла не один год. Далее, в течение более 20 лет, для систематизации и каталогизации библиотечных фондов в Библиотеке ИФП использовались таблицы ББК для научных библиотек [1-5].

В 1997 г. была начата работа по созданию электронного каталога, который изначально был снабжен неплохой поисковой системой и имел немало точек доступа: поиск можно было проводить по автору, заглавию, сведениям, относящимся к заглавию, сведениям об ответственности, параллельному заглавию, заглавию серии, коллективному автору, году издания, редактору (составителю, переводчику и т.п.), персоналиям. В 2009 г. для развития электронного каталога

поисковая система была дополнена предметными рубриками ББК, что позволило проводить поиск литературы аналогично работе с традиционным систематическим каталогом. На рис. 1 представлен перечень отделов и разделов при проведении поиска по шифру. Деление «В379 Физика полупроводников и диэлектриков» выведено на первый уровень, поскольку является профилирующим направлением исследований Института.

На рис. 2 видны подразделы «В379» этого электронного каталога, и вертикальная серая полоска справа показывает, что их примерно половина от общего числа рубрик. Источником рубрик, кроме «В3 Физика», послужили те же таблицы [1-5], базой для «В3» - электронные эталонные таблицы Научно-исследовательского центра (НИЦ) РГБ [6]. Хотя они не имеют алфавитно-предметного указателя, что затрудняет работу по ним, но содержат современные понятия и отражают логику и структуру науки.

Для технической тематики из НИЦ РГБ были запрошены таблицы отделов «Ж Технические науки...»,

«з Радиотехника...» и «К Приборостроение...». И хотя в НИЦ работы с ними еще не были завершены, в 2010 г. условно называемые «промежуточные» рубрики этих отделов были нам предоставлены. Считаем, что впоследствии редактировать их будет проще, чем вариант 60-х гг. XX в. На рис. 3 показаны разделы, подразделы, деления и подразделения «з» в электронном каталоге ИФП СО РАН, на рис. 4 – результат поиска по шифру «з844.15», равный 82 библиографическим записям, включающий как отечественные, так и иностранные издания. Следует обратить внимание, что в результате поиска искомым шифр может стоять в комбинированном индексе как на первом месте, так и на любом другом. Например, в строках 1, 3 (см. рис. 4) он стоит на втором месте. В данном случае книгу по интересующей нас тематике (интегральные схемы микрорадиоэлектронной аппаратуры) в фонде следует искать на полке, маркированной другим шифром, первым в записи (3852 «Полупроводниковые приборы» и 397 «Вычислительная техника» соответственно).

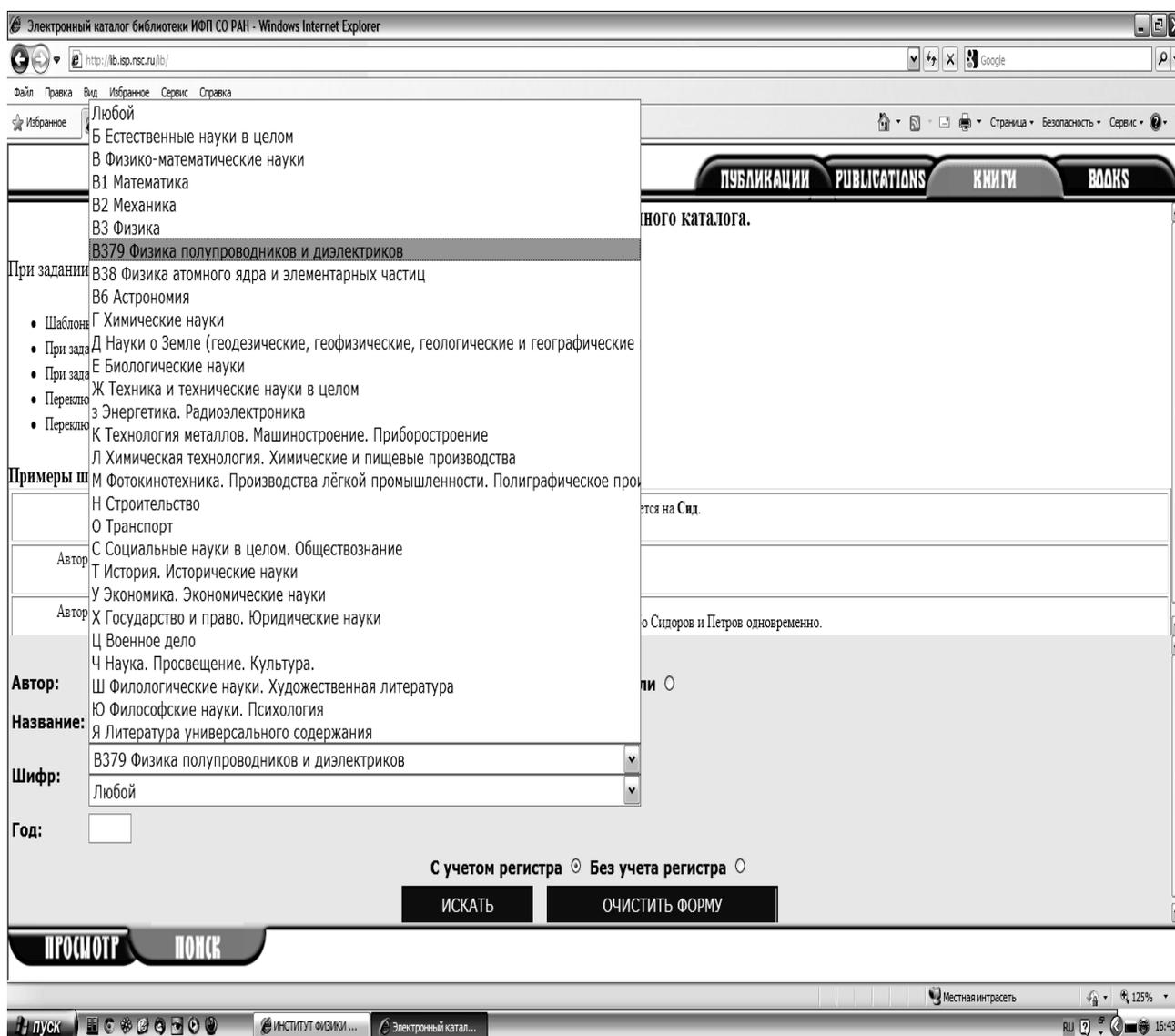


Рис. 1. Предметные рубрики ББК для классификационного поиска литературы в электронном каталоге ИФП СО РАН

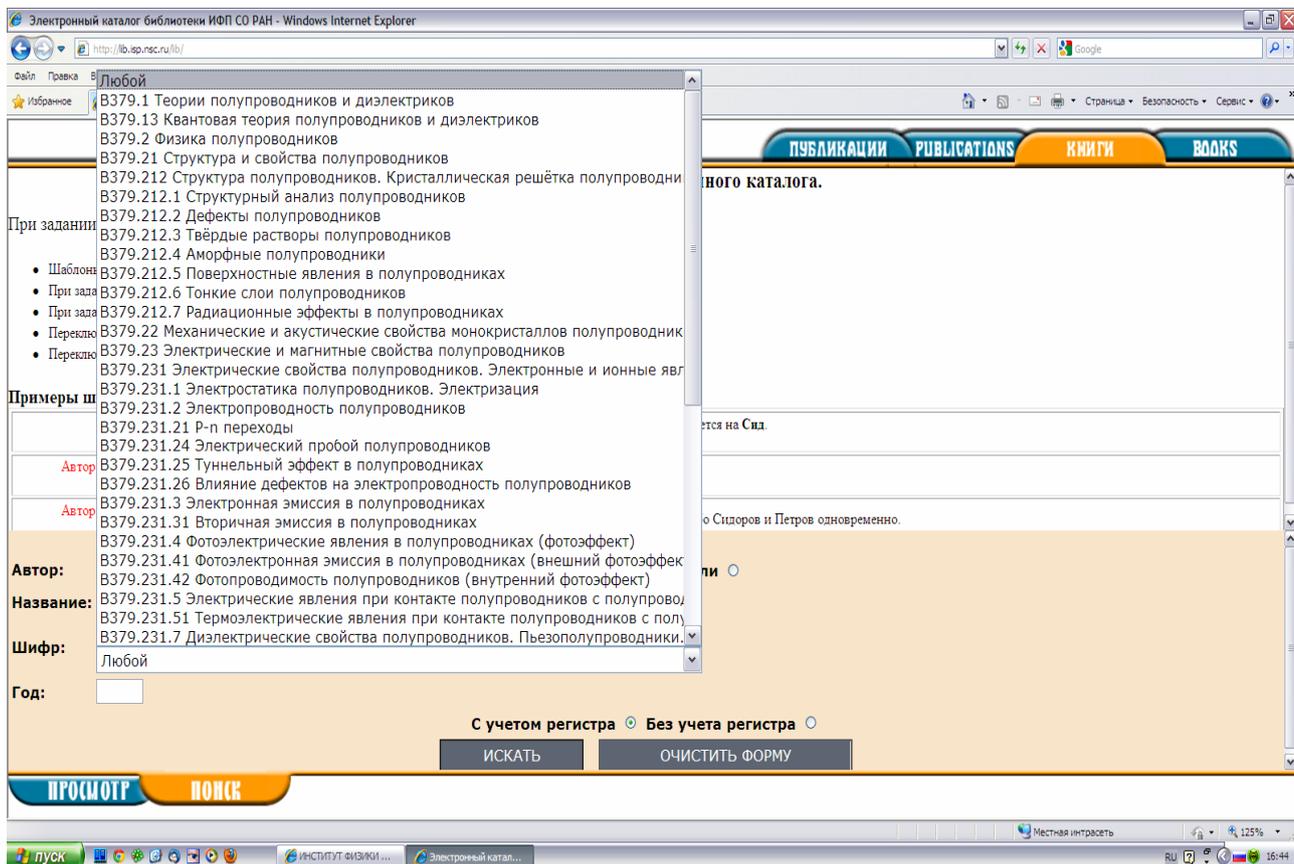


Рис. 2. Подразделения «Физики полупроводников»



Рис. 3. Техническая тематика в электронном каталоге ИФП

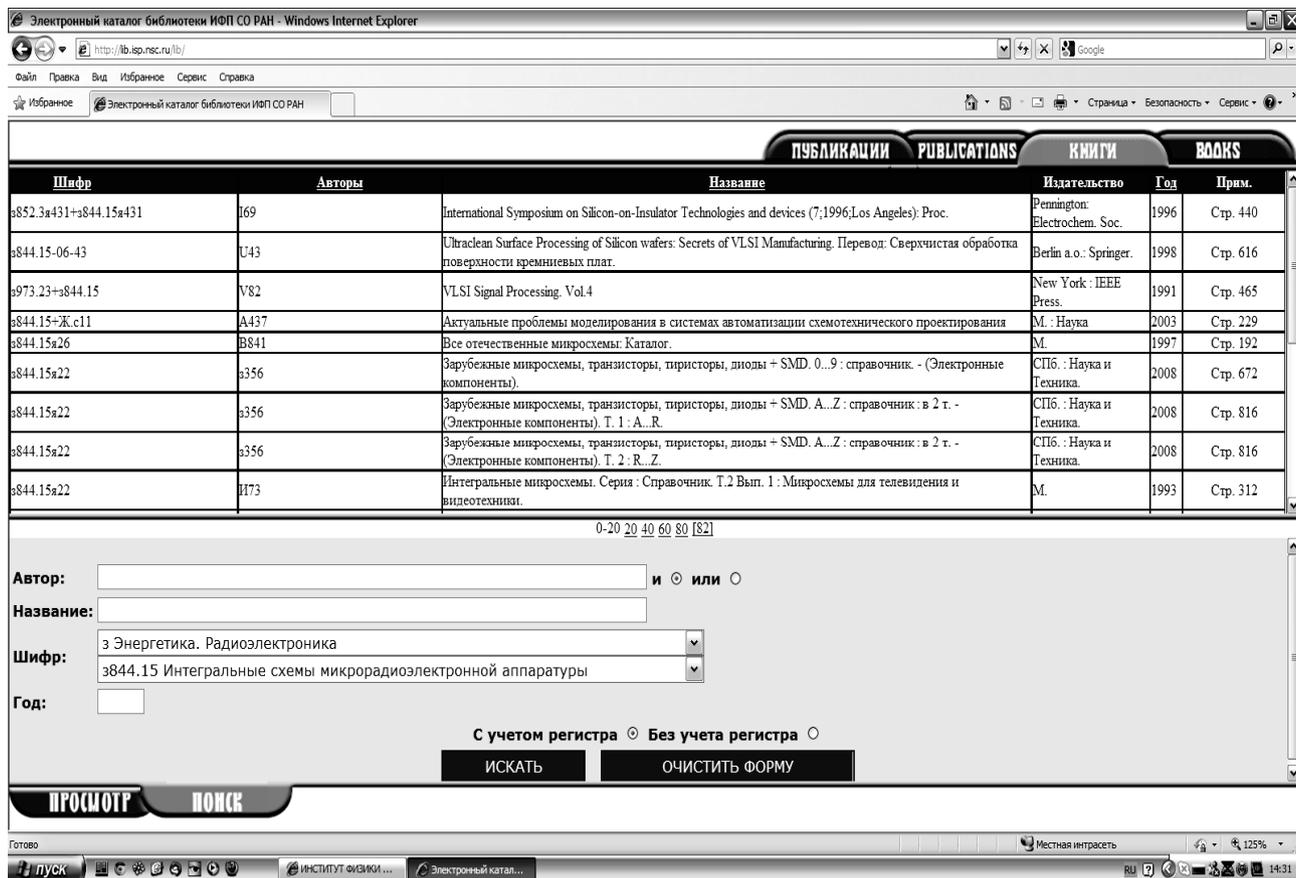


Рис. 4. Результат поиска по шифру «з844.15 Интегральные схемы микрорадиоэлектронной аппаратуры»

В целом, ведение электронного каталога позволило значительно сократить размеры традиционного алфавитного каталога, с 2008 г. в него стали поступать карточки только с основным библиографическим описанием, а возможность проведения классификационного поиска позволила законсервировать в 2012 г. систематический каталог.

Таким образом, можно резюмировать, к чему сводятся проблемы систематизации библиотечных фондов академического института Сибирского региона. Если при переводе фондов и каталогов из УДК в ББК они в большей степени являлись техническими, то в XXI в. они стали касаться содержательной стороны и насыщения таблиц, и принципиальной проблемой стало их отсутствие в переработанном в соответствии с современными тенденциями науки виде.

Так, одной из насущных проблем можно назвать систематизацию нанотематической литературы. Как указано выше, ИФП ведет исследования и научные разработки по элементной базе нанoeлектроники, в том числе физико-химическим основам технологий нанoeлектроники. В состав Института входят лаборатория нанодиагностики и нанолитографии и лаборатория физики и технологии трехмерных наноструктур, сотрудниками которых достигнут ряд практических результатов. Например, к настоящему времени в СО РАН (ИФП, ИЯФ, ИНХ) сформировано 7 новых прецизионных (высокоточных) метаматериалов, которые

невозможно изготовить какой-либо другой известной технологией, кроме нанотехнологий [7].

Поэтому закономерно, что ученые ИФП нуждаются в нанотематической литературе и библиотека активно комплектуется различными изданиями этого профиля. Учитывая, что в настоящее время нанотехнологии признаны приоритетной тематикой на государственном уровне, чем обусловлено обилие потока нанотематической литературы, фонды библиотеки ИФП пополняются ею практически стремительно. Как такую литературу классифицировать? По предложению Э.Р. Сукиасяна логично было бы отнести нанотематику к отделу «1 (А) Общенаучное и междисциплинарное знание (общенаучная и междисциплинарная технология)» [8], потому что нанонаука, нанотехнологии и наноинженерия имеют междисциплинарную основу. Другим вариантом могло бы быть выделение для нанонаук особой рубрики в разделе «Ж Техника и технические науки в целом», поскольку науки приобретают в своём названии приставку «нано», когда они применяются для нанотехнологий и следовательно являются техническими применениями научных результатов [9]. Но пока окончательного решения еще не принято и соответствующих таблиц не существует, внутренним методическим решением Библиотеки ИФП принято, что нанотематическая литература временно собирается в классификационной нише под индексом «з856». Для обеспечения поиска и индексирования этой тематики по другим

классификационным системам используемый нами тезаурус по полупроводникам дополняется рубриками специализированных рубрикаторов нанотехнологий.

## СПИСОК ЛИТЕРАТУРЫ

1. Библиотечно-библиографическая классификация. Таблицы для научных библиотек. Вып. III-VII / М-во культуры РСФСР; ГБЛ. – М., 1960-1964.
2. Библиотечно-библиографическая классификация. Таблицы для научных библиотек. Вып. XXV. Таблицы типовых делений. – М., 1961.
3. Библиотечно-библиографическая классификация. Таблицы для научных библиотек. Сокращенный вариант. Вып. I-V. – М.: Книга, 1970 -1972.
4. Библиотечно-библиографическая классификация. Средние таблицы. Вып. 1-4 / Рос. гос. б-ка [и др.]. – М.: Либерия, 2001-2011.
5. Библиотечно-библиографическая классификация. Средние таблицы. Доп. вып. Таблицы типовых делений общего применения / Рос. Гос. б-ка [и др.]. – М.: Либерия, 2003.
6. Библиотечно-библиографическая классификация. Таблицы для научных библиотек. Раздел «В3 Физика» [Электронный ресурс]. – [Электронная версия эталона таблиц 2009 г. / НИЦ РГБ].
7. О работе Сибирского отделения РАН в 2011 году и задачах на 2012 год : доклад председателя СО РАН академика А.Л. Асеева на Общем собрании Сибирского отделения РАН 19 апреля 2012 г. // Наука в Сибири. – 2012. – № 17. – С.2-5.
8. Сукиасян Э.Р. Междисциплинарное и общенаучное знание: классификационный анализ. Подходы к формированию структуры и содержания нового отдела ББК // Научные и технические библиотеки. – 2010. – № 4. – С. 49-57
9. Белоозеров В.Н., Шабурова Н.Н. Отражение нанотематики в тезаурусе по физике полупроводников // Межрегион. научно-практич. конф. «Научные библиотеки России: взгляд в будущее». Новосибирск, 26-28 сентября 2012 г. – ГПНТБ СО РАН. – Аннотация: URL: [http://conf.nsc.ru/confer\\_2012/ru/reportview/131842](http://conf.nsc.ru/confer_2012/ru/reportview/131842)

УДК 025.4:[002:004]

А.Б. Антопольский (МГУКИ)

## Классификации информационных ресурсов

*Сделан краткий обзор классификаций информационных ресурсов, которые применяются в России для решения задач в разных сферах науки, образования и т.д.*

**Ключевые слова:** *информационные ресурсы, информационные системы, классификации, интеграция информационных ресурсов*

Рост количества и разнообразия информационных ресурсов, доступных через Интернет, поставил заново проблему их классификации. Одной из причин переосмысления задачи классификации является тот факт, что в Интернете представлены категории ресурсов, бывшие ранее объектами различных по назначению информационных систем, и, соответственно, классифицируемых на основе разных теоретических и практических оснований. Причем, это относится к классификациям как традиционных институтов памяти – библиотек, архивов, музеев, органов научной информации, так и других информационных систем: управленческих, торгово-экономических, образовательных, просветительных, массовых и проч.

Наибольший интерес представляют для нас классификации двух категорий: документные классификации и официальные, применяемые для управленческих задач в сфере науки и образования.

Приведем перечень основных документных классификаций, применяемых в России:

- Универсальная десятичная классификация (УДК)

- Библиотечно-библиографическая классификация (ББК)
- Государственный рубрикатор НТИ (ГРНТИ)
- Десятичная классификация Дьюи (ДКД)
- Международная классификация изобретений (МКИ)
- Классификация Архивного фонда РФ
- Классификация государственных стандартов (КГС).

Следует обратить внимание, что эти классификации в Интернете применяются только в тех случаях, когда каталоги на их основе созданы в «доинтернетный» период, иначе говоря, в «офлайне». Крупные порталы, интегрирующие разнородную информацию, документные классификации не применяют. Есть, однако, исключение: портал «Российское образование» использует ГРНТИ, впрочем, наряду с другими классификациями.

Список основных отечественных официальных классификаций, применяемых в сфере науки и обра-

зования, может быть представлен следующими объектами:

- Номенклатура специальностей научных работников ВАК
- Общероссийский классификатор специальностей высшей научной квалификации (ОКСВНК)
- Основные направления фундаментальных исследований
- Классификатор Российского фонда фундаментальных исследований
- Общероссийский классификатор специальностей по образованию (ОКСО)
- Классификация отраслей науки, видов научных исследований, их социально-экономических целей
- Перечни приоритетных направлений развития науки, технологий и техники и критических технологий Российской Федерации.

Эти классификации на практике также не применяются для интеграции разнородной информации в Интернете.

Если говорить о классификациях массового использования, применяющихся для навигации в Интернете, то крупнейшие поисковые машины применяют классификации, специально разработанные для этих целей. В качестве примера приведем первый уровень классификации наиболее популярной в российском Интернете поисковой системы Яндекс:

- Развлечения
- СМИ
- Дом
- Hi-Tech
- Отдых
- Справки
- Работа
- Производство
- Спорт
- Общество
- Учёба
- Авто
- Игровая
- Порталы
- Культура
- Бизнес.

Очевидно, что здесь положен в основу совсем иной принцип, чем в перечисленных выше классификациях и он непригоден для навигации в научно-образовательных ресурсах.

Еще один пример классификации, используемой для навигации в разнородном информационном пространстве – классификация образовательных ресурсов портала ГНИИ ИТТ «Информика»:

- Федеральные образовательные ресурсы
- Региональные образовательные ресурсы
- Учебное книгоиздание и образовательная пресса
- Конференции, выставки, конкурсы, олимпиады
- Инструментальные программные средства
- Электронные библиотеки, словари, энциклопедии
- Ресурсы для администрации и методистов
- Ресурсы для дистанционных форм обучения
- Информационная поддержка ЕГЭ
- Ресурсы для абитуриентов

- Ресурсы по предметам образовательной программы
- Внешкольная и внеклассная деятельность.

В последние годы появились порталы, специально ориентированные на интеграцию информационных ресурсов институтов памяти, т.е. библиотек, архивов, музеев. Целью классификаций, используемых в таких порталах, является навигация, т.е. классификация используется в них наряду с поисковыми машинами и ограниченным перечнем метаанных, заимствованных из каталогов интегрируемых систем. В таких порталах применяются фасетные системы, причем число фасетов варьируется от 3 до 6, не более 3 уровней в каждом.

Приведем два примера. Первый – это логическая структура портала «Культура региона», предложенная коллективом Московского государственного университета культуры и искусств (МГУКИ) с участием автора для Минкультуры России:

- Категории ресурсов
  - публикации
  - архивные документы
  - музейные экспонаты
  - мультимедиа
  - памятники недвижимости
- Тематика
- Персоны
- Организации
- География
- Хронограф
- Словарь.

В основу этой структуры была положена фасетная классификация, разработанная компанией «Альт-софт» для интегрированной тематической системы «Виртуальный музей Гулага»<sup>1</sup>

Второй пример – логическая структура электронной библиотеки «Научное наследие России»<sup>2</sup>. В ней применяется четыре классификационных фасета:

- Категория ресурса
- Тематика
- Хронология
- Территориальный фасет.

При этом фасет «категория ресурса» задан на следующем списке сущностей:

- персоны
- научные объекты (документы, музейные экспонаты, мультимедиа и др.)
- организации
- события
- коллекции.

Очевидно, что фасетная структура этих порталов является переходным этапом к полноценной онтоло-

<sup>1</sup> <http://www.gulagmuseum.org/>

<sup>2</sup> Антопольский А., Каленов Н., Серебряков В., Сотников А. Принципы разработки интегрированной системы для научных библиотек, архивов и музеев // Информационные ресурсы России. – 2012. – № 1. – С. 2-7.

гии, которая теоретически может стать инструментом навигации для универсальной системы поиска научных и образовательных ресурсов в Интернете.

Можно утверждать, что современная ситуация в научно-образовательном информационном пространстве не способствует переходу к практическим ра-

ботам по созданию такой онтологии. Остается, однако, теоретическая постановка вопроса: следует ли стремиться к созданию универсальной поисковой научно-образовательной информационной системы в Интернете или нынешняя ситуация множества разнородных систем является предпочтительной?

УДК 025.4.02[001.103.2:37]

В.А. Старых, А. И. Башмаков (ГНИИ ИТТ «Информика»)

## Классификация и каталогизация информационных ресурсов сферы образования

*Для управления информационными ресурсами (ИР) российской образовательной системы должна быть разработана и принята единая спецификация метаданных, используемая для описания и классификации информации – и в образовательных порталах, и в системе регистрации, и в хранилищах данных. В настоящей работе даётся уточнённое понятие ИР как объекта каталогизации, рассматриваются требования к метаданным, описывающим образовательные ИР, перечислены классификаторы для использования в метаданных и сделаны выводы о принимаемом пути разработки системы метаданных, позволяющей единообразно описывать и классифицировать ИР в сфере образования по их содержанию и пользовательским характеристикам.*

**Ключевые слова:** информационные ресурсы, каталогизация, образовательный портал, метаданные, информационный ресурс образования

### ИНФОРМАЦИОННЫЕ РЕСУРСЫ КАК ОБЪЕКТЫ КАТАЛОГИЗАЦИИ

Выделяются два подхода к определению схем описания ИР в каталогах:

- декомпозиция множества информационных ресурсов (ИР) на классы, для которых создаются разные схемы описания;
- создание единой схемы описания для всех классов ИР с представлением их особенностей с помощью ограничений целостности.

Первый подход предполагает создание множества спецификаций метаданных. Далее возникнет задача их согласования, обобщения, формирования на основе разнородных массивов описаний единых каталогов и т. д.

В настоящее время преобладает тенденция не специализации, а универсализации метаданных, описывающих разные сущности. Поэтому второй подход представляется более обоснованным.

Универсальный характер метаданных не означает, что все виды ИР имеют одинаковый приоритет при наполнении каталогов. Здесь целесообразно установить следующие приоритеты:

- 1) компьютерные ИР с сетевым доступом;
- 2) компьютерные ИР на автономных носителях;

3) прочие виды ИР (включая традиционные бумажные документы).

Разработке универсальной системы классификации и каталогизации ИР будет способствовать принятие строгого определения понятия ИР, которое позволит очертить множество сущностей, подлежащих описанию метаданными, классификации и каталогизации.

Предлагается следующее определение:

Информационный ресурс сферы образования – это совокупность данных, которые могут быть усвоены и преобразованы в знания, организованные для эффективного получения достоверной информации в виде именованного документа, самостоятельно оформленного, предназначенного для распространения среди неограниченного круга лиц и служащего основой для предоставления информационных услуг посредством передачи по информационно-телекоммуникационным сетям и обработки в информационной системе.

Приведенное определение отражает ключевые признаки, которыми должна обладать информационная сущность для того, чтобы быть представленной в классификационной схеме и затем – в общедоступном отраслевом каталоге ИР.

**Первый** ключевой признак определяет, что ИР – это организованная совокупность данных, т.е. имеющая определённую структуру и форму представления, которые предполагают применение технологий обработки с помощью ЭВМ.

**Второй** ключевой признак предполагает, что данные, составляющие ИР, определяют часть знаний, которые формализованы в такой степени, что над ними могут осуществляться процедуры формализованной технологической обработки с помощью различных технических средств.

**Третий** ключевой признак определяет, что ИР – это документ, как объект практической деятельности, с фиксированными реквизитами, позволяющими его идентифицировать<sup>3</sup>, т.е. обладающий некоторой формой представления, обеспечивающей возможность **самостоятельного распространения**. Сущность, не имеющая такого оформления, но зафиксированная с реквизитами в виде информации, позволяющей ее идентифицировать, не может быть отражена в каталоге. Например, абзац из документа и фрагмент файла не распространяются в самостоятельном качестве. В предлагаемой интерпретации к ИР также не относятся интегральные сущности, объединяющие группы ИР со связанными с ними услугами. Например, не является информационным ресурсом учебный курс, включающий услуги преподавателя и образовательного заведения. Необходимо отметить, что этот ключевой признак определяет ИР как именованный объект, более точно – как объект, имеющий некий уникальный идентификатор, обладая которым, ИР наделён следующими качествами:

- 1) благодаря уникальному имени ИР можно отличить от других ИР;
- 2) благодаря уникальному имени к ИР может быть привязано метаописание, позволяющее управлять информацией об этом ИР;
- 3) с помощью уникального имени можно получить доступ к самому ИР.

Необходимо отметить, что команда под руководством Бернерса-Ли предложила решение проблемы именования ИР, которое называется **унифицированный идентификатор ресурса** (uniform resource identifier, URI).

**Четвёртый** ключевой признак определяет ИР как объект, **предназначенный для распространения среди неограниченного круга лиц**, для которых он представляет интерес, т.е. имеет потребительскую ценность. Данная трактовка близка к понятию **публикации** (документа, доступного для массового использования). В то же время ИР – не обязательно опубликованный документ в строгом смысле этого понятия. Размещение информации о нем в каталоге может быть шагом на пути к публикации. Другая обязательная мера для публикации –

обеспечение доступа к ИР, и включение в каталог сведений об условиях доступа (за плату или безвозмездно) неограниченному кругу лиц. Доступ к ИР может быть организован через Интернет или путем выпуска тиража информационных носителей, который удовлетворяет потребности этих лиц и определяется с учетом характера ИР.

Отмеченная особенность позволяет разделить ИР и документы, ориентированные на распространение среди определенного (т.е. ограниченного) круга лиц. К таким документам относятся корпоративные и ведомственные материалы (справки, расписания занятий, учебные планы, ведомости, сертификаты и др.), а также информация для ограниченного использования. Подобные документы остаются за рамками каталогизации.

На рынке интеллектуальной продукции большинство ИР выступают в качестве продуктов или товаров. Исключение составляет особый класс мегаресурсов Интернета, служащих основой для предоставления услуг. Такие ИР не являются товаром (права на их использование не продаются и не покупаются). Один и тот же ИР, как правило, используется для оказания различных услуг. Эти услуги составляют его основную ценность и не могут быть предоставлены без его наличия. Поэтому каталогизации в данном аспекте подлежат ИР, а не услуги. Примеры подобных ИР: электронная библиотека, поисковая система, интернет-каталог, интернет-трансляция и др.

Необходимо подчеркнуть, что предложенная трактовка ИР шире понятия **издание**. Неотъемлемый признак издания – прохождение редакционно-издательской обработки – не является обязательной чертой ИР. Издания на разных типах носителей составляют значительную долю ИР. Однако к ИР также относятся документы, не попадающие в категорию изданий (например, отчет о НИР, депонированная рукопись, диссертация, образовательный объект и др.).

**Пятый** ключевой признак определяет ИР как объект, служащий основой для предоставления информационных услуг.

**Шестой** ключевой признак определяет современную тенденцию сетевой природы распространения ИР посредством передачи по информационно-телекоммуникационным сетям.

**Седьмой** ключевой признак определяет условия создания и использования ИР посредством наличия технологий обработки в информационной системе.

## МЕТАДАННЫЕ ОБ ОБРАЗОВАТЕЛЬНЫХ ИНФОРМАЦИОННЫХ РЕСУРСАХ

Основные требования к схеме метаданных и способу их представления:

- универсальность (в рамках установленного понимания ИР как объекта каталогизации);
- достаточная выразительность для обеспечения эффективного решения задач, требующих наличия метаданных;

<sup>3</sup> Как относительно новое научное явление, электронный документ не получил пока устойчивой и определенной дефиниции.

- совместимость с международными стандартами и протоколами в области метаданных и информационного поиска (создание условий для интероперабельности);

- соответствие российскому законодательству в области образования (в частности, установленным уровням и ступеням образования) и отражение особенностей отечественной образовательной системы;

- возможность задания ограничений целостности, отражающих взаимосвязи полей описания ИР;

- обеспечение возможности хранения метаданных как совместно с ИР, так и отдельно от него.

Основные задачи, при решении которых используются метаданные:

- поиск ИР по метаданным;
- обработка метаданных программными агентами;
- навигация по каталогам ИР;
- формирование покрытия учебного плана образовательными ИР;

- управление контентом, в том числе агрегация ИР (построение интегрального ИР из имеющихся ИР и их компонентов);

- мониторинг состояния ИР образовательной сферы;

- информационно-аналитические исследования.

Подходы к построению схемы метаданных:

- выбор одной из существующих международных схем (Dublin Core, Learning Object Metadata);

- выбор одной из существующих международных схем и ее расширение;

- формирование оригинальной схемы.

В нашей статье рассматривается использование второго подхода.

В качестве основы рассматривается использование схемы метаданных Learning Object Metadata (LOM):

- LOM удовлетворяет почти всем указанным выше требованиям (для отражения особенностей российской системы образования предлагается расширение);

- схема ориентирована именно на образовательные ИР;

- по сравнению с Dublin Core схема LOM более структурирована и обладает большей выразительностью;

- LOM отражает опыт, приобретенный ведущими мировыми разработчиками образовательных ИР и поставщиками образовательных услуг;

- описание ИР на основе LOM является частью манифеста (внешнего формата представления информации, фиксирующей структуру ИР и логику навигации по нему);

- набор обязательных полей, который будет определен в расширении LOM, должен быть достаточно компактным.

## КЛАССИФИКАТОРЫ И СЛОВАРИ В МЕТАДАНЫХ

Предлагаются следующие классификаторы и словари, играющие центральную роль при описании ИР, так как на их основе реализуется большинство ограничений целостности:

- словарь (одноуровневый классификатор) целевых назначений ИР

- классификатор видов ИР (типология ИР).

Классификация ИР по целевому назначению (источники – ГОСТ 7.60-90 и ГОСТ 7.83-2001):

- 1) официальное
- 2) научное
- 3) научно-популярное
- 4) производственно-практическое
- 5) нормативное производственно-практическое
- 6) учебное
- 7) массово-политическое
- 8) справочное
- 9) для досуга
- 10) рекламное
- 11) художественное.

Один и тот же ИР может иметь несколько целевых назначений.

Классификатор видов ИР базируется на типологии видов изданий (ГОСТ 7.60-90).

Базовые основания деления:

- среда, необходимая для использования ИР;
- форма представления информации, содержащейся в ИР, и тип носителя ИР;
- статус ИР как товара на рынке интеллектуальной продукции либо средства, на основе которого оказываются определенные услуги;
- содержание ИР, представляющее основную ценность;
- характер информации, содержащейся в ИР, а также типы решаемых с его помощью задач.

ИР может быть отнесен только к одному виду (листу классификационного дерева).

Классификаторы и словари, которые целесообразно использовать при навигации по каталогу ИР портала:

- словарь целевых назначений ИР
- классификатор видов ИР
- предметный классификатор
- словарь уровней образования
- словарь целевой аудитории (категорий пользователей ИР).

Словарь уровней и ступеней образования – источник Федеральный закон от 29 декабря 2012 г. № 273-ФЗ "Об образовании в Российской Федерации".

Идентификатор вершины	Наименование вершины
1	Дошкольное образование (preschool education)
2	Начальное общее образование (primary education)
3	Основное общее образование (basic education)
4	Среднее (полное) общее образование (secondary education)
5	Начальное профессиональное образование (technical school first cycle)
6	Среднее профессиональное образование (technical school second cycle)
7	Профессиональная подготовка (vocational training)
8	Высшее профессиональное образование (higher education)
8.1	Бакалавриат (bachelor degree)
8.2	Подготовка дипломированных специалистов (graduating)
8.3	Магистратура (magistracy)
9	Дополнительное образование (additional education)
10	Послевузовское профессиональное образование (university postgraduate)

Один и тот же ИР может быть отнесен к нескольким уровням.

Словарь целевой аудитории (категорий пользователей ИР). Этот элемент данных предлагается ассоциировать только с ИР учебного назначения. Он заполняется терминами словаря LOM:

- Teacher – преподаватель
- Author – автор
- Learner – обучаемый
- Manager – менеджер

и дополнительными терминами предлагаемого расширения словаря, которые необходимы для отражения всех категорий пользователей:

- *Methodologist* – методист
- *Instructional Designer* – разработчик образовательных компонентов
- *Graphical Designer* – графический дизайнер
- *Multimedia Developer* – разработчик мультимедийных компонентов

- *Programmer* – программист
- *Technical Implementer* – технический разработчик.

Причины введения расширения словаря и его цели:

- отражение специфики российской системы образования, а также отечественных законодательства и стандартов;

- более глубокая структуризация и формализация ряда элементов описания ИР;
- представление характеристик, специфичных для ИР, распространяемых на информационных носителях;
- развитие ограничений целостности с целью снижения ошибок в метаданных;
- определение профилей описаний для разных классов ИР.

## ВЫВОДЫ

Спецификация метаданных об образовательных ИР строится как расширение схемы LOM.

В качестве базовых классификаторов и словарей, используемых в метаданных, предлагаются:

- словарь целевых назначений ИР (на основе ГОСТ 7.60-90 и 7.83-2001);
- классификатор видов ИР (на основе типологии видов изданий);
- предметные классификаторы (УДК, ГРНТИ);
- словарь уровней образования (в соответствии с Законом РФ об образовании);
- словарь целевой аудитории (расширенный словарь LOM).

Для предложенной информационной модели метаданных разработаны [1, 2]:

- профили описаний ИР разных классов;
- ограничения целостности;
- XML-привязка;
- программные средства формирования метаданных.

## СПИСОК ЛИТЕРАТУРЫ

1. Принципы построения и описания профилей стандартов и спецификаций информационно-образовательных сред: Метаданные для информационно-образовательных ресурсов сферы образования / глав. ред. В. А. Старых; отв. ред. А. И. Башмаков. – М.: Европейский центр по качеству, 2009. – 378 с. – ISBN 978-5-94768-063-8
2. Большаков А. И., Старых В. А. Принципы и технологические основы создания открытых информационно-образовательных сред. – М.: БИНОМ. Лаборатория знаний, 2010. – 720 с. – ISBN 978-5-9963-0285-7.

В. Н. Белоозеров (ВИНИТИ РАН), В. А. Старых (ГНИИ ИТТ «Информика»)

## Онтология классов информационных ресурсов образования

*Заданы критерии описания и классификации предметной области информационных ресурсов образования и построена гамма представлений формальной онтологии области в соответствии с постепенно увеличиваемой глубиной и шириной рассмотрения. В число представлений входят теоретико-множественные структуры, эквивалентные простой номенклатуре, перечислительной классификации и информационно-поисковым тезаурусам.*

**Ключевые слова:** информационные ресурсы образования, онтология предметной области описания, метаданные, метаинформация

Управление в системе информационных ресурсов образования (ИРО) требует наличия формального описания структуры этой системы. Такое описание, основанное на классификации типов информационных ресурсов и спецификации модели их метаданных, предложено в монографии [1] и обосновано в [2]. Однако современные требования к степени формализации описаний предполагают построение **формальной онтологии** предметной области описания. Перед нами стоит задача формального описания системы информационных ресурсов в сфере образовательной деятельности с тем, чтобы вложить представление об этой сфере в автоматизированные системы, помогающие производить создание, учёт, систематизацию, хранение, поиск, преобразование и комбинирование информационных ресурсов с целью организации учебного процесса в образовательных учреждениях различного уровня.

В последнее время такое формализованное описание получило название **онтология**. В переводе на русский язык этот термин означает: *слово о существующем = учение о существе = изложение действительности*. Такое изложение может быть в разных формах. Если изложение ведётся в форме философских рассуждений, то это – **ОНТОЛОГИЯ** как философская дисциплина (**философская онтология**)<sup>1</sup>. Если изложение ведётся в виде инвентаризации имён объектов, свойств объектов, их отношений и происходящих с ними процессов так, что имена можно зафиксировать в виде кодов и алгоритмов компьютерной системы, то такое изложение вполне можно тоже назвать **ОНТОЛОГИЕЙ** – **формальной онтологией**<sup>2</sup>.

<sup>1</sup> Это – исконное традиционное значение слова *онтология*. Когда других значений не существовало, или когда другие значения не актуальны, эпитет «философская» излишен и не употребляется.

<sup>2</sup> Это – новое значение слова *онтология*. Поскольку значительный круг образованных людей до сих пор связывает с этим словом только его философское значение, уточнение понятия эпитетом «формальная» необходимо, чтобы не вызывать у свежего читателя недоумения и отторжения. Здесь мы выполнили требование об уточнении нашего употребления слова

Наша задача – сформировать **формальную онтологию системы информационно-образовательных ресурсов**.

Что касается самой **системы информационно-образовательных ресурсов**, то состав и границы этой **системы** (как и любой другой системы) не имеют чёткой определённости и выявляются в процессе описания, часто **в результате** описания, в зависимости от задачи и аспекта рассмотрения. Вот эту область действительности, ограниченную задачами данного этапа описания, мы будем называть **предметной областью информационных ресурсов образования**. Центральным звеном нашего рассмотрения всегда будет совокупность именно **информационно-образовательных ресурсов**. Рассмотрение других объектов и включение их в рассматриваемую систему зависит от глубины и широты рассмотрения. При **углублении** рассмотрения в предметную область включаются свойства, отношения, процессы, описания, структуры и элементы информационных ресурсов. При **расширении** рассмотрения в предметную область включаются информационные ресурсы иного характера (научные, персональные, юридические, ...), а также образовательные учреждения, обучающий персонал, обучаемый контингент и т. п. Что же касается описания самих информационно-образовательных ресурсов, то его предполагается производить не во всей их полноте, а без учёта индивидуальности конкретных данных о том или ином ресурсе. Описание предполагается вести на уровне типов, характеристик, отличительных черт ресурсов, т. е. на уровне **метаинформации**, не затрагивая вопросов о конкретном содержании (контенте) в информационных ресурсах. Таким образом, в основании нашей предметной области лежат различные **классы информационных ресурсов** сферы образования. Каждый класс является совокупностью информационных ресурсов, наделённых каким-либо общим **свойством**, служащим основанием для вхож-

*онтология*, и дальше мы, как правило, не будем употреблять эпитета «формальная».

дения ресурса в соответствующий класс. Система этих «классообразующих свойств» изоморфна системе самих классов, так что отдельное рассмотрение этих свойств не конструктивно, они «синонимичны» классам, и всё сказанное о классе является сказанным об этих свойствах. Мы же при описании предметной области будем вести речь о свойствах, характеризующих *группы* классов, об отношениях между классами ресурсов, о процессах с классами ресурсов как об отражении свойств, отношений и процессов, имеющих место или возможных в среде конкретных информационных ресурсов.

Поскольку мы во всех случаях будем иметь в виду именно информационно-образовательные ресурсы, мы их будем называть просто *информационные ресурсы* или даже – *ресурсы* и будем обозначать сокращением *ИР*.

Наша задача – построить онтологию предметной области информационных ресурсов сферы образования – будет в определённой мере выполнена тогда, когда:

- будет задан перечень всех (рассматриваемых) имён классов ИР, имён свойств, отношений и процессов,
- для каждого имени класса будут указаны все (рассматриваемые) имена и значения свойств, отношений и процессов, в которые вовлечён класс,
- для каждого имени свойства будут указаны имена всех классов ИР, отношений и процессов, с которыми имеет связь данное свойство с данным значением,
- для каждого имени отношения будут указаны имена всех классов ИР, свойств и процессов, с которыми имеет связь данное отношение,
- для каждого имени процесса будут указаны имена всех классов ИР, свойств и отношений, с которыми имеет связь данный процесс.<sup>3</sup>

При этом построенная онтология будет формальной (теоретикомножественной) структурой  $\Omega = \Omega(C, P, R, S)$ <sup>4</sup>, где  $C$  – словарь имён классов ИР,  $P$  – словарь имён свойств,  $R$  – словарь имён отношений,  $S$  – словарь имён процессов. Здесь каждый словарь состоит из словарных статей, характеризующих свой объект по отношению к объектам других словарей. В частности, каждому заглавному имени в словарной статье должны быть приписаны имена связанных с ним элементов других словарей. В конкретных случаях при ограниченной глубине и широте рассмотрения некоторые из этих словарей могут быть пустыми (могут отсутствовать). Во всех случаях не пуст словарь  $C$ . Если остальные словари пусты  $P = R = S = \emptyset$ , то мы имеем простой перечень классов ИР – **номенклатуру**. Если пусты  $P$  и  $S$ , а  $R$  включает только отношение род-вид, то такая онтология выражается структурой простой иерархической классификации. Если в классификации словарь  $S$  дополнить ограниченным перечнем других отношений и ввести непустой словарь  $P$ , то мы получим структуру развитого

<sup>3</sup> Эти требования к построенной онтологии в дальнейшем будут называться ЦЕЛЬЮ построения.

<sup>4</sup> Латинские и греческие буквы выбраны для обозначения элементов структуры как инициалы следующих английских слов: Class, Property, Relation, Sentence (утверждение о процессе) и греческого  $\Omega v$  (существующий).

тезауруса, соответствующего стандартам в области информационно-поисковых языков (ГОСТ 7.25 [3], ISO 5964 [4])<sup>5</sup>.

По мере углубления и расширения описания в состав предметной области могут войти и другие объекты кроме указанных выше.

**1. Номенклатура классов.** Простейшее описание сущего – это перечисление сущностей, что называется **номенклатурой** объектов предметной области. Состав объектов, подлежащих обсуждению, мы будем определять на основе монографии [1], в которой содержится наиболее полное и достаточно формализованное описание системы информационных ресурсов образования через типологию их метаданных. Задача настоящего исследования состоит в том, чтобы представить структуру изложенных сведений в едином формате, допускающем непосредственное использование в различных автоматизированных системах на требуемом уровне глубины и широты описания.

Перечень классов информационных ресурсов образования задан в монографии [1] классификационной таблицей (Таблица 1, с. 23 – 29). Перечень ИР представляет только одну из составляющих содержания таблицы. Он может быть алфавитным списком, приведённым на сайте <http://systemling.narod.ru/informat/OntoIRE/Content.doc> (далее – Приложение)<sup>6</sup>. Этот список и предстаёт как онтология первого (поверхностного) уровня, соответствующая поставленной ЦЕЛИ:

$$\Omega_1 = \Omega(C_1, \emptyset, \emptyset, \emptyset), \quad (1)$$

где  $C_1$  – список классов ИР, а  $\emptyset$  – знаки пустых множеств свойств, отношений и процессов.

**2. Синонимия.** Оставаясь на уровне перечисления классов ИР, мы можем увидеть, что имена многих классов предполагают изменчивость. Так, форма имени «Презентация / Демонстрация» предполагает, что данный класс ИР может именоваться также именами «Презентация» и «Демонстрация». А имя «Тезисы доклада (сообщения) на научной конференции (съезде, симпозиуме)» допускает изменение даже двух слов, и не все варианты включены в  $C_1$ , а общее количество форм этого имени – более 10. Таким образом, классы ИР именуется не одним именем, а некоторыми множествами равнозначных имён. В теории информационного поиска такие множества называют термином «дескриптор», а между отдельными формами дескриптора считают установленным

<sup>5</sup> Стандарты на информационно-поисковые тезаурусы впервые были разработаны в 60-х гг. прошлого века, когда концепция формальных онтологий для сетевых информационных ресурсов ещё не могла возникнуть. В настоящее время принята новая версия международного стандарта, которая учитывает современные потребности [5].

<sup>6</sup> Онтология информационных ресурсов образования со всеми необходимыми для использования таблицами и программными средствами будет доступна на сайте <http://spec.edu.ru>, где она будет служить компонентом интеллектуальной системы управления ресурсами. В настоящее время с таблицами онтологии можно ознакомиться по адресу: <http://systemling.narod.ru/informat/OntoIRE/Content.doc>. Авторы выражают благодарность к. т. н. Ю. П. Косарской за помощь в формировании таблиц.

отношение синонимии, или эквивалентности.<sup>7</sup> При таком подходе словарь имён должен быть расширен, и для каждой формы дескриптора должен быть указан список синонимичных форм. Этот словарь  $C_2$  реализует более глубокое представление онтологии информационных ресурсов образования. Учитывая вариативность именования классов ИР, получаем формальную онтологию:

$$\Omega_2 = \Omega(C_2, \emptyset, R_2, \emptyset), \quad (2)$$

где  $R_2$  – множество отношений между элементами словаря  $C_2$ , включающее всего один элемент – **отношение синонимии**, которое обозначим греческими буквами  $\Sigma$  (от слова  $\Sigma\nu\nu\nu\alpha$  = равноимённость).  $R_2 = \{ \Sigma \}$ .

Отношение синонимии очевидным образом разбивает словарь  $C_2$ , на систему непересекающихся подмножеств, каждому из которых сопоставлен один элемент словаря  $C_1$ . Эту формальную систему можно поэтому представить также и без введения отношения  $\Sigma$ , а именно как онтологию:

$$\Omega_3 = \Omega(C_2, P_3, \emptyset, \emptyset), \quad (3)$$

где  $P_3$ , - **свойство** элементов словаря принадлежать к классу синонимов<sup>8</sup> для того или иного элемента словаря  $C_1$ . Сами элементы  $C_1$  могут выступать как **значения** этого свойства. Иначе говоря  $P_3 : P_3(C_2) = C_1$ .

Таким образом, онтология классов ИР с учётом вариативности имён может быть представлена двумя способами, которые согласуются с поставленной выше ЦЕЛЬЮ:

- перечнем всех форм именования классов ИР (перечнем элементов словаря  $C_2$ ), для которых указаны «все» отношения (т. е. одно отношение  $\Sigma$ ), и перечнем «всех отношений», для которых (для единственного отношения  $\Sigma$ ) указаны все элементы  $C_2$ , вступающие в отношения (т. е. указаны все пары синонимичных наименований классов).

- перечнем всех элементов словаря  $C_2$ , для которых указано свойство быть возможной формулировкой имени класса, указанного в  $C_1$ , и перечнем всех классов (т. е. элементов множества  $C_1$ ), для которых указаны все элементы множества  $C_2$ , наделённые этим свойством.

Последний вариант представлен в Приложении (Таблица В) по сути дела всего лишь одной таблицей, представленной в двух видах – Таблица В1 и Таблица В2. Эти таблицы состоят из двух столбцов. В одном столбце содержится перечень элементов словаря  $C_1$ , а в другом – соответствующих им элементов  $C_2$ . Разница заключается только в расположении столбцов и в упорядоченности элементов в них. Второй способ более глубоко отражает онтологию ИР, поскольку указывает, какой из синонимов дескриптора является предпочтительным именем класса.

Здесь можно таблицы В1 и В2 слить в одну, упорядочить по алфавиту и ликвидировать дублирование имен дескрипторов. В получившейся Таблице В3 перечислены все дескрипторы, для которых указаны возможные вариации их имени (называемые «аскрипторами», или «синонимами дескриптора»), а также перечислены все аскрипторы (*показаны курсивом*) с указанием на соответствующий дескриптор. Такая структура в практике информационного поиска носит наименование «**дескрипторный словарь**». В таблице В3 отношение синонимии распадается на два взаимнообратных отношения – С: («синоним», отражение дескриптора на совокупность его синонимов) и См: («смотри», отражение аскриптора на соответствующий синонимичный дескриптор). Обозначения С: и См: выбраны в соответствии с ГОСТ 7.25 как сокращения слов «синоним» и «смотри» соответственно<sup>9</sup>. Эти сокращения мы будем использовать как для обозначения ссылок между дескриптором и аскрипторами в таблицах элементов онтологии, так и для обозначения соответствующих формальных отношений между элементами. С использованием отношений С: и См: онтологию классов ИР можно представить структурой  $\Omega_{3a}$ , которая содержательно эквивалентна структуре  $\Omega_3$ .

$$\Omega_{3a} = \Omega(C_2, \emptyset, R_{3a}, \emptyset), \quad (3a)$$

где  $R_{3a} = \{ С:, См: \}$  – множество отношений, состоящее из двух указанных элементов. Отношение С: – многозначное, оно определено во множестве дескрипторов с областью значений, составляющей множество всех аскрипторов, т. е. у дескриптора может быть несколько синонимов или ни одного. Отношение См: – однозначное, оно определено на множестве всех аскрипторов и отображает его во множество дескрипторов, т. е. у каждого аскриптора имеется один и только один дескриптор, являющийся стандартным именем соответствующего класса ИР, а дескрипторы разных аскрипторов могут совпадать.

**3. Иерархия классов.** Назначение Таблицы 1 в монографии [1] – не просто перечислить классы ИР, но задать классификацию их типов и назначений. Это – более глубокий уровень описания предметной области. В перечень типов ИР вводится отношение взаимного включения классов. Точнее, вводится два связанных отношения, которые обозначим русскими буквами ВР: и НВ:<sup>10</sup>. Отношение ВР: – *выше-род* – сопоставляет некоторому имени класса ИР  $x$  имя охватывающего класса  $y$ :  $ВР(x) = y$  если  $y \supset x$ . Это же самое обстоятельство выражается через отношение *ниже-вид*:  $НВ(y) = x$ . Формальная онтология ИР приобретает теперь вид

$$\Omega_4 = \Omega(C_2, P_3, R_4, \emptyset), \text{ где } R_4 = \{ ВР:, НВ: \} \quad (4)$$

<sup>7</sup> Точнее, дескриптором называют как класс синонимичных имён объектов, так и одно из этих имён, выбранное для представления класса в целом.

<sup>8</sup> На содержательном уровне свойство синонимии означает, что классы ИР, обозначаемые различными терминами-синонимами, тождественны.

<sup>9</sup> Установленные ГОСТом обозначения отношений здесь и далее мы дополнили знаком двоеточия для более яркого отражения функций этих сокращений. В англоязычных тезаурусах соответствующие отношения обозначают символами *UF (used for)* и *USE* – согласно международному стандарту ISO 2788.

<sup>10</sup> Эти обозначения читаются так: *ВР* – *выше-род*, *НВ* – *ниже-вид*. Они следуют стандарту на представление информационно-поисковых тезаурусов ГОСТ 7.25. В англоязычных тезаурусах им соответствуют обозначения *BTG* и *NTG (broader term generic, narrower term generic)*.

В Таблице 1 отношение ВР/НВ выражено путём присвоения именам классов ИР (дескрипторам имён классов) иерархических числовых кодов, указывающих на взаимное включение классов друг в друга. В совокупности с расположением списка имён классов согласно возрастанию кодовых номеров получается классификационная таблица, которая наглядно и компактно представляет структуру онтологии ИР на данном уровне рассмотрения. Эту же структуру можно представить согласно поставленной нами ЦЕЛИ. Для этого следует в перечне дескрипторов (имён классов ИР) каждому дескриптору приписать имена всех классов, связанных с данным классом отношениями ВР: и НВ:, а для каждого из этих отношений указать все пары имён классов ИР, связанных данным отношением. Соответствующее дополнение таблицы ВЗ превращает её в форму, предусмотренную для **лексико-семантического указателя тезауруса** согласно ГОСТ 7.25 (см. таблицу Г1 в Приложении), где для дескрипторов указаны имеющиеся подклассы с пометой НВ: и надклассы с пометой ВР:. При этом сама Таблица 1 (её левая колонка) может рассматриваться как **иерархический указатель**, требуемый для представления тезауруса стандартом ГОСТ 7.25.

Таблица Г1 совместно с левой колонкой Таблицы 1 из монографии [1] образуют адекватное нашей ЦЕЛИ представление онтологии ИР в варианте

$$\Omega_5 = \Omega(C_2, \emptyset, R_5, \emptyset), \text{ где } R_5 = \{C:, \text{См.}, \text{ВР.}, \text{НВ.}\} \quad (5)$$

**4. Полииерархия.** Ещё на этапе составления словарей  $C_1$  и  $C_2$ , мы обратили внимание на то, что большая часть классов ИР распадается на две категории в соответствии с типом носителя данных – компьютерные и бумажные. И для различения «одноимённых» ресурсов компьютерные представления мы снабдили релятором {1}, а бумажные – релятором {4} в соответствии с номером раздела Таблицы 1. При этом для каждого бумажного ИР находится соответствующий компьютерный, имеющий идентичное наименование и содержание (контент). Этот факт формально никак не отражён в Таблице 1, но имеет большое значение для пользователей ИР, желающих получить информацию независимо от того, на каком носителе она представлена. Онтология в структуре информационно-поискового тезауруса по ГОСТ 7.25 в состоянии отразить дополнительные связи дескрипторов такого типа. Для этого достаточно ввести в словари  $C_1$  и  $C_2$  наименования соответствующих классов без пометок о разделе Таблицы 1. Эти дескрипторы будут отражать классы ИР, объединённые контентом, независимо от носителя, а их представления на том или ином носителе будут соответствующими подклассами, связанными родовидовыми отношениями ВР: и НВ:. Эти связи будут показаны в дескрипторных статьях наряду с иерархическими связями левого столбца Таблицы 1.

При этом окажется, что ряд дескрипторов входят одновременно в разные родовые объединения, в разные иерархические ряды. Например:

Автореферат диссертации	НВ: Автореферат диссертации {1} НВ: Автореферат диссертации {4}
Автореферат диссертации {1}	ВР: Автореферат диссертации ВР: Электронные представления бумажных изданий и документов
Автореферат диссертации {4}	ВР: Автореферат диссертации ВР: ИР на бумажных носителях

Такие полииерархические связи не удаётся отразить обычными классификационными таблицами, но они естественно входят в тезаурусную форму представления онтологий.

С учётом этих отношений мы получаем более глубокое представление реалий ИР в форме онтологии

$$\Omega_6 = \Omega(C_3, \emptyset, R_5, \emptyset), \text{ где } R_5 = \{C:, \text{См.}, \text{ВР.}, \text{НВ.}\} \quad (6)$$

Здесь  $R_5$  формально не претерпевает изменения, но содержательно отношения ВР:, НВ: расширились за счёт включения новых пар дескрипторов, связанных этими отношениями. Лексико-семантический указатель тезауруса, соответствующего онтологии  $\Omega_6$  приведён в **Приложении Д**. Иерархический указатель тезауруса следует дополнить отдельным перечнем вновь введённых обобщающих дескрипторов со ссылками на их виды в соответствии с носителем данных. Этот перечень уже содержится внутри лексико-семантического указателя.

Кроме видов ИР Таблица 1 монографии [1] задаёт список «назначений» ИР. Термины назначений можно считать свойствами классов ИР. Они образуют множество  $P_4$ . В свою очередь классы ИР, объединённые общим для них назначением (свойством из  $P_4$ ), нетрудно обозначить соответствующими именами, составленными из имени класса и имени назначения, например

Вид ИР – *базы данных*

Назначение – *научное*

Имя класса – *научные базы данных.*

Такие составные имена следует образовать для обозначения классов ИР, которым свойственно то или иное назначение в соответствии с отметкой плюсом (+) в Таблице 1. А каждому из назначений в целом соответствует свой класс ИР, обозначаемый одним из следующих одиннадцати имён

*официальные ИР*

*научные ИР*

*научно-популярные ИР*

*производственно-практические ИР*

*нормативно-практические ИР*

*учебные ИР*

*массово-политические ИР*

**справочные ИР  
ИР для досуга  
рекламные ИР  
художественные ИР.**

Эти новые имена классов ИР оказываются связанными друг с другом и с именами из словаря  $C_3$  теми же самыми родовидовыми отношениями ВР и НВ. При этом новые «комбинированные» имена классов (с указанием назначения) будут наделены полииерархическими связями: с одной стороны с соответствующим именем ИР без указания назначения, а с другой стороны – с именем обобщающего класса ресурсов данного назначения. Например:

научные базы данных	ВР: базы данных
	ВР: научные ИР

В соответствии со ссылками на родовые классы, для имён классов без указания назначения могут быть перечислены соответствующие видовые классы (в именах которых указано назначение), например

базы данных	НВ: научные базы данных
	НВ: производственно-практические базы данных
	НВ: нормативные производственно-практические базы данных
	НВ: учебные базы данных
	НВ: справочные базы данных
	НВ: рекламные базы данных

Для каждого обобщающего класса назначения информационных ресурсов могут быть перечислены конкретные классы такого назначения, например:

официальные ИР	НВ: официальный электронный журнал
	НВ: официальная электронная газета
	НВ: официальная Internet-публикация
	НВ: официальная газетная/журнальная публикация {1}
	НВ: официальное уставное издание {1}
	НВ: официальная инструкция {1}
	НВ: официальный массив инструкций
	НВ: официальный стандарт {1}
	НВ: официальный массив стандартов
	НВ: официальный бюллетень {1}
	НВ: официальный массив бюллетеней
	и т. д.

С учётом всего этого мы получаем полное описание исходной Таблицы 1 в виде онтологии

$$\Omega_7 = \Omega(C_4, \emptyset, R_5, \emptyset),$$

где

$$R_5 = \{C_3, \text{См.}, \text{ВР.}, \text{НВ.}\}, \quad (7)$$

а  $C_4$  – это  $C_3$ , дополненный выше указанными именами.

Эта онтология, так же как и предыдущая  $\Omega_6$ , адекватно представляется лексико-семантическим указателем тезауруса с полииерархией. Таблица этого указателя будет представлена в Приложении Е. Поскольку это довольно большое дополнение, работа будет выполняться по мере осознания её практической целесообразности.

**5. Ассоциация.** Типичный информационно-поисковый тезаурус согласно ГОСТ 7.25 имеет в своём составе также «ассоциативные» связи дескрипторов. Они устанавливаются между объектами, связанными друг с другом в процессах своего жизненного цикла так, что информация по одному из них может оказаться полезной для интересующихся другим объектом. Обычным источником таких связей являются «ассоциации по смежности» (часть-целое, сырьё-продукт, причина-следствие и т. п.) и «ассоциации по сходству» (совпадение внутренней структуры объектов).

В Таблице 1 также можно найти основания для установления ассоциативных связей. Так, например, в таблице рядышком стоят следующие наименования ИР:

*автореферат диссертации {4}* и  
*массив авторефератов диссертаций*  
*инструкция {4}* и  
*массив инструкций*  
*стандарт {4}*  
и *массив стандартов*  
*научная работа, статья из сборника* и  
*сборник научных работ, статей*  
и т. п.

Это типичные ассоциации по смежности типа часть-целое. Ассоциации по сходству могут быть установлены, например, между дескрипторами *отчёт о НИР* и *отчёт об УНИР*, *магистерская диссертация* и *дипломный проект/работа*.

Указанные ассоциативные связи введены в лексико-семантический указатель тезауруса, представленный в Приложении таблицей Д. Ассоциации по смежности обозначены ссылками АСМ., ассоциации по сходству – АСХ.:

В итоге мы здесь получаем тезаурусное представление онтологии ИР на уровне левого столбца Таблицы 1, которая использует все возможности, рекомендуемые по ГОСТ 7.25, и даже несколько углубляет описание онтологии ИР по сравнению с Таблицей 1.

**6. Категории дескрипторов.** Назначение информационных ресурсов может быть отражено в тезаурусной форме более экономным образом, чем предложено в разделе 4. Для этого нужно ввести в словник тезауруса  $C_3$  сами термины назначений

вместо комбинаций наименований классов ИР с их возможными назначениями. При этом новый словник  $C_5$  по сравнению с  $C_3$  возрастает всего на одиннадцать терминов вместо многочисленных наименований ИР с указанием назначения (и сами термины короче). Эти термины можно взять в форме, указанной в Таблице 1:

- официальное*
- научное*
- научно-популярное*
- производственно-практическое*
- нормативное производственно-практическое*
- учебное*
- массово-политическое*
- справочное*
- для досуга*
- рекламное*
- художественное*

Эти 11 дескрипторов составляют в  $C_5$  особую семантическую категорию, поскольку обозначают не классы ИР, а некоторые атрибуты классов. Это выражено грамматической формой терминов: ИР выражены именами существительными (точнее – номинативными фразами), а назначения – прилагательными (атрибутивными фразами). Формализовать такое различие дескрипторов можно введением в онтологию двух свойств: «*обозначает класс ИР*» и «*обозначает назначение ИР*», которые можно указать в дескрипторных статьях специальными пометами, предусмотренным в стандартах на тезаурус, например пометами «*К: ИР*» и «*К: назначение*».<sup>11</sup> А между классами ИР и возможными назначениями должна быть указана связь, которая по смыслу является ассоциацией по смежности, но в силу специфики этой связи мы обозначим её особыми символами: *Нз:* в статьях дескрипторов категории «*ИР*» и символом *Зн:* в статьях категории «*назначение*».

Пример.

Тезаурусная статья дескриптора категории «*класс информационных ресурсов*»

База данных	К: ИР
	ВР: Информационные продукты
	Нз: научное
	Нз: производственно-практическое
	Нз: нормативно-производственно-практическое
	Нз: учебное
	Нз: справочное
	Нз: рекламное

Тезаурусная статья дескриптора категории «*назначение информационных ресурсов*»

научное	К: назначение
	Зн: База данных
	Зн: Презентация/демонстрация
	Зн: Электронный журнал
	Зн: Мультимедийная запись
	и т. д.

Такое представление соответствует онтологии

$$\Omega_8 = \Omega(C_5, P_4, R_6, \emptyset),$$

где

$$R_6 = \{C:, C_m:, ВР:, НВ:, АСМ:, АСХ:, Нз:, Зн:\} \quad (8)$$

$C_5$  – это  $C_3$ , дополненный одиннадцатью именами назначений

$$P_4 = \{К:ИР, К:назначение\}.$$

Лексико-семантический указатель тезауруса с учётом вновь введённых дескрипторов и связей находится в Приложении – таблица Ё. Он адекватно ЦЕЛИ представляет Таблицу 1 монографии [1] и онтологию  $\Omega_8$ .

**7. Уровни образования и целевая аудитория.** Информационные ресурсы каждого класса, обозначаемого дескрипторами построенной Таблицы Ё1 (первая часть таблицы Ё), могут различаться по другим характеристикам, важными для системы образования. Такими характеристиками являются: (1) соответствие материала ИР ступени образовательного процесса (уровню образования) и (2) предназначение лицам, играющим разную роль в образовательном процессе. Приведём списки соответствующих понятий согласно монографии [1]:

**Уровни образования:**

- дошкольное
- начальное общее
- основное общее
- среднее (полное) общее
- начальное профессиональное
- среднее профессиональное
- высшее профессиональное
- послевузовское профессиональное
- профессиональная подготовка
- дополнительное

**Целевая аудитория:**

- преподаватели
- авторы
- обучаемые
- менеджеры
- методисты
- исследователи
- абитуриенты
- разработчики образовательных ресурсов
- графические дизайнеры
- разработчики мультимедийных компонентов
- программисты
- технические разработчики.

Эти характеристики, также как и «назначение», могут быть связаны с классами ИР. Однако в монографии [1] эти данные не представлены. Поэтому мы от-

<sup>11</sup> Эти обозначения следует читать так: К: ИР = «категорией дескриптора является класс ИР» и К: назначение = «категорией дескриптора является назначение ИР».

ложим этот аспект на этап последующих доработок. В настоящее время вышеприведенные списки уровней образования и целевых аудиторий можно рассматривать как отдельные «номенклатурные» онтологии, дополнительные к основной онтологии, представленной таблицей Ё1. В принципе эти списки можно влить в таблицу Ё1, снабдив каждый термин пометой «К: уровень» и «К: аудитория» соответственно, но пока связи этих терминов с дескрипторами других категорий не установлены, делать это нет смысла.

Согласно монографии [1] таблицу Ё1 можно рассматривать как онтологию, соответствующую классификации информационных ресурсов образования

## СПИСОК ЛИТЕРАТУРЫ

1. Большаков А. И., Старых В. А. Систематизация информационных ресурсов для сферы образования: Классификация и метаданные / Россий-

ский гос. университет инновационных технологий и предпринимательства. – М.: Европейский центр по качеству, 2003. – 384 с. – ISBN 5-94768-002-5.

2. Старых В. А., Башмаков А. И. Классификация и каталогизация информационных ресурсов сферы образования // НТИ. Сер. 1. - 2013. - № 4. – С. 27-30
3. ГОСТ 7.25-2001. Тезаурус информационно-поисковый одноязычный. Правила разработки, структура, состав и форма применения. – М.: Стандартинформ, 2001. – 16 с.
4. ISO 2788:1986. Documentation – Guidelines for the establishment and development of monolingual thesauri. – Genève: ISO, 1986. – 35 p.
5. ISO 25964-1:2011. Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval. – Genève: ISO, 2011. – 152 p.

УДК 001.5:[94(=16):025.45.05УДК]

Митрович Божидар

## УДК: Исследования истории славян до VI века\*

*Рассматривается отражение в УДК непрерывности и связи по времени культур на территории Европы. Утверждается, что территориальное разделение народов не отражает их исторические генетические связи. В частности, нецелесообразно географическое разделение славян на восточных, южных и западных. Географически разделённые лужицкие сорбы (в Германии) и балканские сербы объединяются с восточными славянами исконным исторически правильным термином руси/русины/русские, который не отражён в УДК. Не соответствует историческому прошлому размещение старославянского языка в классе южнославянских языков, так как это был язык всех славян. Не учитывает УДК и результаты исследований, доказавших возможность прочтения этрусских и некоторых других древних текстов на основе славянских соответствий [1, 2].*

*Единство культуры и истоков науки европейских народов обосновывается нами языковыми соответствиями терминов, имеющих латинское происхождение, с терминами и представлениями древних славян о мироустройстве. Можно утверждать, что забытие и сокрытие таких соответствий подобно терминологической войне против славянства. Носителями и распространителями единой культуры европейских народов мы считаем представителей особой расовой группы, идентифицируемой неуничтожимыми маркерами в геноме. Эти маркеры обнаруживаются на территории Балканского полуострова 12 тысячелетий тому назад и сегодня распространены по территории всех индоевропейских народов Европы и Азии. Представителей этой группы можно называть «протославянами», равно как и предками других европейских народов. На основе этих соображений мы предлагаем пересмотреть классификацию народов в УДК с тем, чтобы можно было продолжить историю славянства в глубь тысячелетий за пределы VI в. новой эры.*

**Ключевые слова:** УДК, исследование истории славян

## СПИСОК ЛИТЕРАТУРЫ

1. Светислав С. Билбия Староевропски језик и писмо Етрураца . – Београд: «Мирослав», 2000.

2. Божидар Трифунов Митрович. КолоВени (СлоВени) и континуитет културе и права. – Београд: КолоВения, 2008. – 756 с.

\* На научно-практической конференции автор сделал полный доклад по указанной теме. Оргкомитет принял решение поместить в сборнике только реферат статьи.

О. А. Антошкова, В. Н. Белоозеров Е. Ю. Дмитриева (ВИНИТИ РАН)

## Стандарты, используемые в информационной деятельности

*Сделан обзор работ по стандартизации объектов, продуктов и процессов информационной деятельности, включая задачи научно-информационного обслуживания, библиотечно-библиографической работы, делопроизводства, издательского и архивного дела. Показаны основные направления унификации и оптимизации методов представления, поиска и обеспечения потребителей содержательной информацией. Сформулированы задачи и ближайшие перспективы развития системы информационных стандартов.*

**Ключевые слова:** научно-техническая информация, библиотечное и издательское дело, СИБИД, Технический комитет по стандартизации, информационная деятельность

Миссия ВИНТИ РАН как головного органа научной и технической информации состоит не только в предоставлении данных потребителям, но и в том, чтобы содействовать оптимизации структуры информационных ресурсов, средств поиска и получения из них научных знаний. На эту цель направлены работы по стандартизации объектов, продуктов и процессов информационной деятельности, которые ведутся в рамках функционирующего при ВИНТИ РАН (Научно-технологическое отделение) Технического комитета по стандартизации ТК 191 «Научно-техническая информация, библиотечная и издательская деятельность». ТК 191 организует разработку национальных, межгосударственных и международных стандартов, образующих специализированную подсистему в рамках отечественной системы стандартизации – «Систему стандартов по информации, библиотечному и издательскому делу (СИБИД)».

Основные цели национальной стандартизации в Российской Федерации определены ГОСТ Р 1.0\*. Для системы СИБИД они конкретизируются в ГОСТ Р 7.0.0-2010\*\* тем, что стандарты системы СИБИД устанавливают единые оптимальные нормы информационных процессов и обеспечивают:

- информационную совместимость;
- конкурентоспособность информационной продукции, работ и услуг за счёт нормализации информационных процессов на передовом уровне международных стандартов;
- конкурентоспособность в других областях деятельности за счёт эффективного информационного обеспечения разработки, производства и реализации продукции, работ и услуг;

- ускорение научно-технического прогресса при эффективном информационном обеспечении;
- рациональное использование ресурсов за счёт предоставления информации о ресурсах и имеющихся резервах;
- совместимость и взаимозаменяемость технических средств;
- сопоставимость результатов исследований и измерений технических и экономико-статистических данных;
- сравнительный анализ характеристик продукции;
- информационную поддержку внедрения инноваций, выполнения поставок и государственных заказов, арбитражных и судебных решений.

Стандарты системы СИБИД способствуют повышению уровня безопасности во всех сферах жизни за счёт обеспечения беспрепятственного распространения объективной научной и технической информации. Самостоятельной целью системы СИБИД является повышение уровня информационной безопасности.

Главная цель системы СИБИД – создание нормативной базы для формирования общегосударственного распределённого фонда научных, технических, образовательных и социальных информационных ресурсов, включая:

- обеспечение качества информационной продукции и услуг;
- создание средств каталогизации, классификации и индексирования информации;
- развитие системы поиска и представления научных, технических и образовательных данных;
- обеспечение сохранности и оптимального управления информационными ресурсами.

Унификация нормативов информационных процессов стандартами системы СИБИД проводится с целью внедрения в практику оптимальных решений по типовым задачам информационной деятельности, исходя из опыта отечественных информационных

\* ГОСТ Р 1.0- 2004 Стандартизация в Российской Федерации. Основные положения. – М: ИПК Издательство стандартов, 2005.

\*\* ГОСТ Р 7.0.0-2010 Система стандартов по информации, библиотечному и издательскому делу. Основные положения. – М: Стандартинформ, 2010.

органов, библиотек, издательств и архивов, а также из данных международной стандартизации.

ТК 191 — это объединение специалистов, являющихся полномочными представителями заинтересованных организаций — информационных и документационных центров, библиотек, издательств, архивов. Возглавляет ТК 191 — секретариат, ведение которого возложено на Научно-технологическое отделение ВИНТИ РАН. В составе комитета работают подкомитеты по следующим направлениям:

- научно-техническая информация;
- библиотечное дело;
- издательское дело;
- управление документацией.

Всего силами ТК 191 разработано и поддерживается более 100 национальных и межгосударственных стандартов России и ведётся работа более чем по 160 международным стандартам ИСО, которые относятся к тематике технического комитета ИСО/ТК 46.

Тематически стандарты СИБИД можно разбить на следующие группы:

- терминология информационной деятельности,
- информационные процессы и инструменты,
- библиотечно-библиографическая деятельность,
- издательская деятельность,
- управление и сохранение документов.

Стандарты на термины и определения задают систему понятий и единые способы выражения знаний в области информационной деятельности, что обеспечивает единство представлений у всех участников информационных процессов и тем самым способствует их эффективному взаимодействию.

Стандарты на информационные процессы и инструменты охватывают всю цепочку обработки документальной информации и включают как интеллектуальные «ручные» процессы, так и автоматическую обработку данных:

аналитико-синтегическую переобработку документов (аннотирование, реферирование, обзорные процессы);

кодирование информации (кодовое представление текстов, письменностей, условных знаков, численных данных, информационных ресурсов);

форматы представления данных, обеспечивающие информационную совместимость участников обмена данными, единство их интерпретации;

средства индексирования и поиска необходимой информации — информационно-поисковые языки (ключевые слова, рубрикаторы, классификации знания, информационно-поисковые тезаурусы).

Стандарты издательской деятельности обеспечивают формирование первичных документов информационного процесса в форме, оптимальной для их функционирования в обществе, включая наличие в документах данных, необходимых для последующих процессов распространения информации. Они содержат требования к оформлению традиционных и электронных изданий, выходные сведения, метаданные, стандартную нумерацию и статистический учёт информационных ресурсов.

Стандарты библиотечно-библиографической деятельности определяют требования к первичной об-

работке изданных документов (традиционных и электронных), к организации их хранения и распространения. Сюда входят правила библиографического описания, каталогизации, систематизации, предметизации и библиотечной статистики.

Стандарты по управлению документами определяют работу с неизданными информационными материалами в организациях и архивах с целью обеспечения их сохранности и доступности для пользователя.

В качестве примера важнейших стандартов, которыми информационные работники должны пользоваться буквально каждый день, можно привести следующие:

ГОСТ 7.1 СИБИД. Библиографическая запись. Библиографическое описание. Общие требования и правила составления

ГОСТ 7.9 СИБИД. Реферат и аннотация. Общие требования

ГОСТ 7.14 СИБИД. Формат для обмена информацией. Структура записи

ГОСТ 7.23 СИБИД. Издания информационные. Структура и оформление

ГОСТ 7.25 СИБИД. Тезаурус информационно-поисковый одноязычный. Правила разработки, структура, состав и форма представления

ГОСТ 7.32 СИБИД. Отчёт о научно-исследовательской работе. Структура и правила оформления

ГОСТ 7.50 СИБИД. Консервация документов. Общие требования

ГОСТ 7.54 СИБИД. Представление численных данных о свойствах веществ и материалов в научно-технических документах. Общие требования

ГОСТ 7.56 СИБИД. Издания. Международная стандартная нумерация сериальных изданий

ГОСТ 7.59 СИБИД. Индексирование документов. Общие требования к систематизации и предметизации

ГОСТ 7.70 СИБИД. Описание баз данных и машинночитаемых информационных массивов. Состав и обозначение характеристик

ГОСТ 7.82 СИБИД. Библиографическая запись. Библиографическое описание электронных ресурсов. Общие требования и правила составления

ГОСТ 7.83 СИБИД. Электронные издания. Основные виды и выходные сведения

ГОСТ 7.90 СИБИД. Универсальная десятичная классификация. Структура, правила ведения и индексирования

ГОСТ Р 7.0.4 СИБИД. Издания. Выходные сведения. Общие требования и правила оформления

ГОСТ Р 7.0.49 СИБИД. Государственный рубрикатор научно-технической информации. Структура, правила использования и ведения

ГОСТ Р 7.0.53 СИБИД. Международный стандартный книжный номер. Использование и издательское оформление

ГОСТ Р 7.0.66-2010 (ИСО 5963:1985) СИБИД. Индексирование документов. Общие требования к координатному индексированию

ГОСТ Р ИСО 15489-1-2007 СИБИД. Управление документами. Общие требования.

В настоящее время система стандартов продолжает развиваться главным образом в направлении обеспечения единства и качества автоматизированных процессов. На национальном уровне находятся в разработке и представлены для обсуждения проекты национальных стандартов по темам:

- Электронные издания. Виды и основные требования
- Коды для представления наименований письменностей
- Упрощённая транслитерация русского письма
- Терминология делопроизводства.

На международном уровне ведётся работа по обсуждению следующих проектов стандартов ИСО:

- Международная стандартная ссылка на документ

- Международный стандартный код звукозаписи
- Применение радиочастотных меток в библиотеках
- Пересмотр ИСО 25577 MarcXchange (Перекодировка Marc – XML)
- Запись китайского текста латинскими буквами
- Терминология по информации и документации
- Международная музейная статистика
- Пересмотр ИСО 8 Оформление периодических изданий.

Приглашаем заинтересованных специалистов высказать свои замечания по предлагаемым проектам. Это будет способствовать более полному учёту интересов российских информационных служб и повышению качества разрабатываемых стандартов.