

РОССИЙСКИЙ ЭТАЛОН УДК: ВЕДЕНИЕ И СОВЕРШЕНСТВОВАНИЕ

*Белоозеров В.Н., Антошкова О.А.,
Астахова Т.С., Смирнова О.В.
(ВИНИТИ РАН)*

1. Универсальная десятичная классификация (УДК) является одним из самых авторитетных и используемых средств упорядочения информационных ресурсов в мире. Приблизительная оценка свидетельствует, что по УДК систематизировано около трети мировых книжных фондов. В нашей стране УДК используется многими издательствами, публикующими материалы по науке и технике, библиотеками и информационными центрами для индексирования печатной продукции и других ресурсов научно-технической информации. Успешное функционирование этой всеобщей и универсальной схемы знаний сопряжено с необходимостью проведения определённых работ по обеспечению этого функционирования. Необходимо, с одной стороны, обеспечить соответствие схемы внутренней структуре современного знания, которое постоянно развивается и видоизменяется, а с другой стороны, необходимо обеспечить соответствие потребностям пользователей, применяющих УДК в информационных процессах на основе современной развивающейся технологии.

Эта работа ведётся в рамках достаточно развитой инфраструктуры. Владельцем интеллектуальных прав и методическим центром УДК является международный консорциум издателей таблиц — UDC Consortium со штаб-квартирой в Гааге [1]. Консорциум определяет правила использования классификационной системы в информационных процессах, ведёт эталонный массив классификационных таблиц на английском языке — так называемый UDC Master File, вносит в него необходимые изменения, распространяет по подписчикам эталонные таблицы и бюллетени изменений к ним. Консорциум также определяет политику издания и ведения иноязычных таблиц УДК — на отдельных национальных языках. В

этом году Консорциум открыл свободный доступ через Интернет к многоязычной базе данных верхних уровней УДК — UDC Summary, которая включает около 2-3 тысяч наиболее важных классификационных позиций. Таким образом, краткая версия таблиц УДК доступна на 43 языках мира. Русская часть этого информационного ресурса подготовлена усилиями коллектива ВИНИТИ.

2. ВИНИТИ в лице директора академика Ю.М. Арского входит в состав управляющего Консорциумом Исполнительного комитета, и Консорциум делегировал ВИНИТИ осуществление своих прав и обязанностей относительно издания таблиц УДК на русском языке. В осуществление этой миссии мы выполняем следующие работы:

- завершено четвёртое полное издание классификационных таблиц (10 томов, или 12 книг);
- регулярно издаются бюллетени изменений и дополнений к изданным таблицам;
- подготовлено переиздание первого тома с учётом всех изменений и дополнений (этим мы открываем практику специальных изданий таблиц УДК по разделам, претерпевшим значительные изменения.);
- выпущено 4 издания кратких таблиц (на уровне 4 тысяч позиций);
- совместно с заинтересованными организациями разработаны специализированные отраслевые таблицы УДК (по книжному делу, информатике, сельскому хозяйству, информационно-образовательным ресурсам);
- разработаны и изданы учебные пособия пользователям УДК;
- опыт работы ВИНИТИ сведён в специальную монографию, которая вышла из печати в конце 2010 г. [2];
- издания ВИНИТИ по УДК предоставляются пользователям в бумажном и электронном виде;
- разрабатывается система навигации по информационным ресурсам на основе сопоставительных таблиц классификационных систем, где УДК занимает центральное место;

- регулярно проводятся семинары для повышения квалификации пользователей УДК;
- работает консультационная служба по сложным вопросам индексирования по УДК, дающая ответы на запросы по телефону, электронной почте и на специальном сайте [3];
- при ВИНТИ сформирован Межрегиональный методический совет пользователей УДК как форум для обсуждения и принятия решений по вопросам эффективного использования классификационной системы.

Вместе с ВИНТИ в поддержке УДК участвуют другие заинтересованные организации, в частности, ГПНТБ России и БЕН РАН.

ГПНТБ России оказывает пользователям УДК ценную услугу, разработав систему автоматизированного рабочего места индексатора по УДК и наладив распространение его по библиотекам страны.

Всё это позволяет индексаторам и систематизаторам эффективно использовать существующие таблицы УДК для систематизации и описания содержания научной и технической документации. Нашими изданиями пользуются не только в России, но в странах ближнего зарубежья, включая Украину и Прибалтику, несмотря на то, что там разрабатывают таблицы УДК на своих национальных языках.

3. Уже исполнилось более 100 лет, как УДК была разработана и стала применяться на практике. За это время в её структуре произошло не так много изменений, но за век бурного развития знаний содержание классов не могло не измениться, порой кардинально. Со времён последнего советского издания в УДК произошли существенные изменения; вот наиболее заметные, которые практически исключают возможность пользоваться соответствующими разделами старых «советских» изданий:

- пересмотрены и перенесены в другое место разделы вычислительной техники, метрологии, менеджмента;

- изменены обозначения языков, религий, исторических периодов, вопросы социального страхования, физической химии;
- исключены общие определители аспектов и введены определители свойств, отношений и процессов;
- введён класс нанотехнологий;
- произведены крупные частные изменения во многих классах (экономика, транспорт и др.);
- вслед за политическими изменениями всё время вносятся изменения в таблицу территориальных определителей.

Эти изменения внесены Консорциумом УДК, главным образом, на основе собственных разработок без достаточного согласования с практическими пользователями системы в библиотеках. В результате возникли определённые трудности, в частности, в связи с игнорированием «правила 10 лет», когда новое содержание оказывалось вложенным в старые коды с необходимостью срочно переиндексировать крупные документальные фонды. Однако новый главный редактор УДК в Консорциуме Анда Славич налаживает взаимодействие с «низовыми» работниками. При Консорциуме образован консультативный совет экспертов из числа непосредственных исполнителей, представляющих разные страны. Мы (ВИНИТИ), со своей стороны, стараемся привлечь непосредственных пользователей русскоязычными таблицами УДК во вновь образованный методический совет, чтобы представлять в Консорциум согласованное мнение наших пользователей.

Несмотря на отмеченные выше изменения, некоторые другие разделы УДК явно требуют дальнейшего совершенствования. К таковым относится раздел физики, который несёт на себе следы представлений даже не прошлого, а позапрошлого века. Но здесь мы имеем некоторый успех: недавно Консорциум по нашему представлению принял и опубликовал изменения к разделу, разработанные Л.А. Верной, сотрудницей БЕН РАН. Но проблемы здесь ещё остаются.

Раздел математики, разработанный в своё время сотрудниками ВИНИТИ, также долгое время не знал дополнений и стал недостаточным для точного индексирования

узкоспециальных статей. Значительное углубление соответствующих разделов было разработано группой математиков под руководством И.Ю. Никольской. Эти дополнения также были приняты и опубликованы Консорциумом.

Совершенно не соответствует современным потребностям класс 007, отведённый для вопросов кибернетики. Предложения по его совершенствованию представлены на данном семинаре Н.П. Шапшевой, которой приходится индексировать соответствующие материалы в БЕН РАН [4].

И, как часто бывает, сапожник — без сапог. В УДК нет достойного места для классифицирования вопросов информатики, информационных технологий, научно-технической информации. Есть вычислительная техника, но аспекты содержательного, смыслового рассмотрения информационных процессов здесь выглядят чужеродными. Есть библиотечное дело, где рассмотрены вопросы формального описания содержания документов, но очень слабо представлены процедуры поиска информации. И совсем не представлены процедуры аналитико-синтетической обработки документов в целом и не предусмотрены вопросы автоматизации. Есть, наконец, класс 002.6 «Документационные центры» (в русском переводе — «Органы НТИ»), но здесь представлен только аспект территориальных различий: межнациональные — национальные — корпоративные органы.

Давайте обратимся к опыту других авторитетных классификаций — американской Десятичной классификации Дьюи (ДКД) и российской Библиотечно-библиографической классификации (ББК). Классификация Дьюи является фактическим стандартом для межбиблиотечного обмена данными, а ББК обязательна для индексирования всей отечественной книжной продукции. И в ДКД (в русском переводе) [5] и в ББК (в новом томе «средних таблиц») [6] все функции библиотек и информационных центров объединены в одном разделе под примерно одинаковым наименованием «Библиотечная и информационная деятельность». По этому же пути следует пойти и в отношении раздела УДК 02, расширив его содержание так, чтобы он включал как традиционные библиотечные

технологии, так и процессы аналитико-синтетической обработки и автоматизированного поиска информации, которые преимущественно присущи именно центрам научно-технической документации. Мы в своё время уже разработали соответствующее предложение [7]. Теперь этот проект мы предполагаем обсудить и доработать в рамках Межрегионального методического совета для представления на утверждение Консорциуму УДК.

К этим кардинальным вопросам совершенствования содержания системы добавляются и текущие изменения в наименованиях географических объектов в связи с объединением ряда субъектов РФ и исправлением неточностей. Так что на сегодня мы ведём разработку следующие вопросы, которые предлагаем для обсуждения:

- Пересмотр класса *007 Кибернетика*;
- Совершенствование класса *53 Физика*;
- Развитие класса *51 Математика*;
- Отражение вопросов информатики;
- Корректировка определителей стран и некоторых других общих определителей;
- Разработка методики поиска по индексам УДК.

4. До настоящего времени совершенствование УДК заключалось в согласовании классификационной таблицы с естественным содержанием проблем современного научного познания. Но сейчас изменились и условия применения классификационных систем. Жизнеспособность УДК напрямую зависит от того, насколько она окажется полезной в условиях глобальносетевой информатики, в Интернете. Здесь положение УДК оставляет желать лучшего. В немалой степени это обусловлено тем, что наличие индекса УДК на документе ещё не гарантирует того, что этот документ будет найден по той теме, которая зафиксирована в индексе. Для такой гарантии необходимо зафиксировать и узаконить чёткие правила поиска по индексам УДК. Как ни странно, такие правила отсутствуют в имеющихся методических материалах. Имевшийся в своё время (в относительно далёком прошлом) отрицательный опыт применения УДК в автоматизированных системах обусловлен именно отсутствием таких правил, а также слабостью тогдашней информационной техноло-

гии. Развитие технологий привело к возможности решения многих информационных задач без привлечения сложных классификационных систем, но оно же привело к такому усложнению самого информационного пространства, что эффективное ориентирование в нём стало невозможным без моделирования этого пространства внутри информационной системы. Отражением этого обстоятельства стала концепция «семантического веба», где поиск информации осуществляется на основе «онтологии», отражающей структуру предметной области.

Универсальная десятичная классификация и есть вариант такой «онтологии», универсальный, проверенный многолетней практикой, внедрённый в реальные информационные процессы. Можно с разных сторон критиковать УДК, но лучше иметь требуемую онтологию, чем изобретать велосипед заново. Разработка стандартной технологии поиска по индексам УДК и сделает её онтологией всего универсума знаний (пусть онтологией недостаточно развитой, но реально существующей). Наличие стандартных процедур поиска приведёт к тому, что на них будут ориентироваться индексаторы, что уменьшит имеющийся разбой индексирования.

Второе, что необходимо, это установить практику всестороннего (многоаспектного) индексирования документа, при котором в индексе были бы отражены все существенные темы документа, по которым разумна выдача документа на запрос. Такой многоаспектный индекс по необходимости будет комбинированным из простых кодов отдельных классов УДК. Требуемая методика поиска должна определять процедуру анализа сложных индексов и процедуру ранжирования выдачи на основе оценки сходства индекса документа с индексом запроса. Разработка такой методики откроет перспективу применения УДК для поиска в сетях, когда ищущее лицо не может ориентироваться на специфическую традицию формирования индексов в конкретном информационном фонде. Общепринятая стандартная методика поиска по индексам УДК будет существенным совершенствованием системы, и мы планируем работать в этом направлении.

Методика поиска по комбинированным индексам основывается на специфической алгебре индексов УДК,

разработка и анализ которой может быть объектом серьёзных математических исследований. А реализация методики в программах информационных систем будет задачей программистов. Но первое слово тут принадлежит информационным работникам, которые должны разработать методику на идейном уровне. Именно в этом направлении нам видится основная перспектива необходимого совершенствования УДК, а также сходной с ней отечественной ББК.

Литература

1. Сайт Консорциума: <http://udcc.org>.
2. Индексирование фундаментальных научных направлений кодами информационных классификаций: Универсальная десятичная классификация. // Антошкова О. А., Астахова Т. С., Белоозеров В. Н., Смирнова О. В., Соловьёва И. М., Сукиасян Э. Р., Сурикова Н. Г., Чумакова Н. Ф — М.: ВИНТИ, 2010. — 322 с. — ISBN 978-5-94577-046-1.
3. <http://forum.udcc.ru/>.
4. О совершенствовании раздела УДК 007 смотрите отдельную публикацию в настоящем сборнике.
5. Десятичная классификация Дьюи и Относительный указатель: Перевод с английского. 21-е изд. — Тома 1 — 2.. / Рук. Я. Л. Шрайберг, отв. ред. Е. М. Зайцева.— М.: ГПНТБ России, 2000.
6. Сукиасян Э. Р. «Библиотечная и информационная деятельность». Новый раздел Средних таблиц Библиотечно-библиографической классификации. // 14-я международная конференция и выставка LIBCOM-2010. Пансионат «Ершово», 15-19 ноября 2010 г. «Информационные технологии, компьютерные системы и издательская продукция для библиотек». — Доступ: <http://www.gpntb.ru/libcom10/disk/24.pdf>.
7. Белоозеров В. Н. Информационные классификации и электронные библиотеки // Информационные ресурсы России, Вып. 2 (96), 2007. — С. 19-25. (См. также: <http://systemling.narod.ru/udc/LibrariesInUdc.htm>).