

Сравнительный анализ электронных версий классификационных систем

Comparative analysis of e-versions of classification systems

Ю. В. Смирнов

*Государственная публичная научно-техническая библиотека России,
Москва, Россия*

Yury Smirnov

*Russian National Public Library for Science and Technology,
Moscow, Russia*

В статье проводится анализ различных видов электронных версий классификационных систем, наиболее распространенных в России. Автор дает их краткое описание и обращает особое внимание на их лицензионную чистоту и частоту обновлений. В конце статьи дается сравнительная таблица всех электронных версий классификационных систем.

Ключевые слова: классификационные системы, электронные версии, УДК, ББК, ГРНТИ.

The author analyzes different e-versions of classification systems popular in Russia. He describes them and characterizing their license FTO and update rate. He also includes the comparison table for the classifications electronic versions.

Keywords: classification system, e-version, UDC, BBK [Russian Library Bibliographic Classification], GRNTI.

Каждая библиотека в своей деятельности использует одну или несколько классификационных систем. В России наиболее популярными являются:

- Библиотечно-библиографическая классификация (ББК);
- Универсальная десятичная классификация (УДК);
- Государственный рубрикатор научно-технической информации (ГРНТИ).

Обновленные таблицы, а также изменения и дополнения, изначально выпускаются в печатном виде, однако при всей нашей любви к бумажной книге, такие издания обладают некоторыми недостатками, к которым можно отнести:

1. отсутствие периодичности выпусков;
2. накопление большого количества изменений и дополнений, что увеличивает объем томов, в которых происходит поиск нужного понятия (например, для 4-го издания УДК, опубликованного ВИНТИ РАН в 2001-2009 гг., выпущено уже шесть выпусков изменений и дополнений, т.е. на сегодняшний день таблицы УДК насчитывают 16 томов (10 томов базового издания и 6 томов изменений и дополнений));
3. отсутствие у некоторых классификаций единого алфавитно-предметного указателя (АПУ), который призван облегчить поиск необходимого индекса (например, средние таблицы ББК).

Часть этих недостатков могут решить электронные версии классификационных систем. На сегодняшний день нами было замечено 7 видов таких версий:

1. программа, содержащая электронный вариант текста таблиц УДК и ББК;
2. различные интернет-сайты с таблицами классификаций;
3. сканированные копии печатных таблиц;
4. файл в формате Microsoft Word, содержащий таблицы УДК с метками полей;
5. UDC Summary;
6. UDC Online;
7. базы данных (БД) классификационных таблиц УДК, ББК, ГРНТИ, ОКСО, выпускаемые ГПНТБ России и Ассоциацией ЭБНИТ.

Все вышеперечисленные электронные версии обладают разными возможностями, а также различаются лицензионной чистотой. Остановимся на каждой из них подробнее.

Программа, содержащая электронный вариант текста таблиц УДК и ББК. Обладает встроенным простым поиском по последовательности символов. Авторство этой программы установить не удалось, также неизвестна ее лицензионная чистота и качество обновлений. Несмотря на заявленное наличие таблиц УДК, они по факту отсутствуют. Также не указано какие таблицы ББК (средние, сокращенные, для массовых библиотек или для детских и школьных библиотек) представлены в данной программе.

Различные интернет-сайты с таблицами классификаций. Лицензионная чистота и качество обновлений также неизвестны. Чаще всего обладают простым посимвольным поиском и представляют таблицы в виде списка гиперссылок. Иногда таблицы могут быть представлены в виде иерархического дерева, либо другим, удобным для создателя конкретного сайта, способом.

Сканированные копии печатных таблиц. Лицензионная чистота неизвестна. Обычно обновляются в соответствии с печатными выпусками, что иногда вносит неразбериху, поскольку все обновленные версии независимо от года выпуска помечаются словами «новое издание» или «последняя версия».

Если после процесса сканирования была использована технология распознавания текста, то такие таблицы обладают простым посимвольным поиском, встроенным в программу для просмотра файлов PDF. Если же эта технология не применялась, то в такой версии поиск невозможен.

Файл в формате Microsoft Word, содержащий таблицы УДК с метками полей [1], выпускается ВИНТИ РАН, правообладателем эталонных таблиц УДК на русском языке. Обновляется в соответствии с печатными выпусками.

Метки полей представлены в угловых скобках, например:

<01> – иерархический код класса (индекс УДК);

<02> – наименование класса;

<05> – методические указания;

<07> – указание на подразделения по аналогии (подразделять как...);

<09> – ссылки на другие классы и т.д.

В качестве примера приведем запись одного индекса в данном формате:

<01> 615.479.5

<02> Одежда для пациентов

<07> 615.479.5 подразделять как 615.479.4

<09> 615.477 Ортопедический материал. Резиновые изделия и т. п.

Такой файл обладает встроенным в текстовый редактор поиском, который по нашему мнению, не очень удобен, поскольку затрудняет не только составление поискового запроса, но и восприятие результатов поиска.

Однако данный файл может оказаться полезным организациям, создающим собственную автоматизированную библиотечно-информационную систему (АБИС).

UDC Summary [2] доступна на интернет-сайте консорциума УДК, содержит основные деления УДК (2600 классов УДК), а не полные таблицы. Поиск отсутствует, однако есть навигация по иерархическому «дереву» классификации. Обновляется в соответствии с эталонными таблицами на английском языке.

Доступ к данному продукту предоставляется бесплатно. Эта версия является полностью лицензионной и распространяется по лицензии Creative Commons Attribution Share Alike 3.0, которая позволяет копировать, распространять, адаптировать и использовать в своих продуктах, даже на коммерческой основе.

Основные деления УДК представлены в переводе (полном или частичном) на 56 языков. На сегодняшний день перевод на русский язык является частичным.

UDC Online [3] также находится на сайте консорциума УДК, но содержит полные таблицы УДК. Обновляется в соответствии с эталонными таблицами на английском языке. Имеет 8 языковых версий, однако версия на русском языке отсутствует. Имеется расширенный поиск, в том числе по исключенным индексам. Навигация осуществляется по иерархическому «дереву».

Доступ к данному продукту, в отличие от UDC Summary, является платным.

Базы данных (БД) классификационных таблиц УДК, ББК, ГРНТИ, ОКСО, выпускаемые ГПНТБ России и Ассоциацией ЭБНИТ [4] по договорам с правообладателями. Обновляются в соответствии с печатными выпусками.

Поиск в базах данных осуществляется как по индексам, так и по ключевым словам, включая развитую логику запросов и возможность усечения слов. Для БД таблиц УДК реализован поиск в том числе и по исключенным индексам. Навигация может осуществляться как по иерархическому дереву, так и по ссылочному аппарату.

В этих БД присутствует очень полезная и удобная функция для индексаторов – отбор и построение сложных и составных индексов.

На сегодняшний день предлагается шесть таких баз данных:

- 1) полные таблицы УДК;
- 2) сокращенные таблицы УДК;
- 3) средние таблицы ББК;
- 4) таблицы ББК для детских и школьных библиотек;
- 5) рубрикатор ГРНТИ;
- 6) общероссийский классификатор специальностей по образованию (ОКСО).

Все предлагаемые базы данных включают в себя последние изменения и дополнения и полностью соответствуют печатным изданиям. Также следует отметить, что каждая БД является «полным собранием сочинений» соответствующих таблиц классификации.

Данные БД имеют несколько версий: самостоятельный продукт на CD-ROM и встроенный компонент в систему автоматизации библиотек ИРБИС 64 (САБ ИРБИС 64). В качестве самостоятельного продукта на CD-ROM эти БД могут рассматриваться как рабочее место систематизатора. В качестве встроенного компонента в САБ ИРБИС 64 они могут использоваться:

- 1) в автоматизированном рабочем месте каталогизатора (АРМ Каталогизатор), увеличивая скорость поиска и составления индекса по сравнению с использованием печатных изданий;
- 2) в АРМ для читателя (АРМ Читатель) и интернет-версии САБ ИРБИС 64 (АРМ Web-ИРБИС), предоставляя возможности поиска необходимого индекса и навигации по иерархическому «дереву» классификации.

Ознакомившись со всеми электронными версиями классификационных систем, наиболее распространенных в России, можно построить сравнительную таблицу.

	1	2	3	4	5	6	7
УДК	-	+	+	+	+	+	+
ББК	?	+	+	-	-	-	+
ГРНТИ	-	+	+	-	-	-	+
Другие классификационные системы	-	+	+	-	-	-	+ (ОКСО)
Лицензионность	?	?	?	+	+	+	+
Обновления	?	?	?	+	+	+	+
Поиск	Посимвольный	Посимвольный	Посимвольный или отсутствует	Встроенный в программу просмотра файлов	Отсутствует	Расширенный	Расширенный
Навигация	-	+	-	-	+	+	+
Русификация	+	+	+	+	Частично	-	+
Построение сложных индексов	-	-	-	-	-	+	+
Системные требования	ОС MS Windows	Любой веб-браузер	Программа для просмотра нужного типа файлов	Любой текстовый редактор с поддержкой формата “.doc”	Любой веб-браузер	Любой веб-браузер	ОС MS Windows

	1	2	3	4	5	6	7
Использование в других продуктах	-	-	-	-	-	-	+ (САБ ИРБИС 64)
Адаптирование под собственные нужды	-	-	-	+	+	-	По договору

Цифрами в таблице обозначены виды электронных версий классификационных систем:

1. программа, содержащая электронный вариант текста таблиц УДК и ББК;
2. различные интернет-сайты с таблицами классификаций;
3. сканированные копии печатных таблиц;
4. файл в формате Microsoft Word, содержащий таблицы УДК с метками полей;
5. UDC Summary;
6. UDC Online;
7. базы данных классификационных таблиц УДК, ББК, ГРНТИ, ОКСО, выпускаемые ГПНТБ России и Ассоциацией ЭБНИТ.

Согласно вышеприведенной таблице можно сделать следующие выводы:

1. использование первых трех электронных версий (№№ 1 – 3) является нарушением авторских прав составителей и разработчиков классификационных систем;
2. лицензионные версии таблиц ББК поставляются только в виде БД, выпускаемых Ассоциацией ЭБНИТ (№ 7) совместно с РГБ;
3. лицензионные версии таблиц ГРНТИ поставляются только в виде БД, выпускаемых ГПНТБ России (№ 7) совместно с ВИНТИ РАН;
4. лицензионные версии УДК, поставляемые составителями и разработчиками (№№ 5 – 6), имеют с точки зрения русскоязычного пользователя следующие недостатки (например, отсутствие русификации (№ 6), частичная русификация (№ 5));
5. несмотря на такие преимущества, как бесплатность и лицензионность, версия (№ 5) обладает таким недостатком, как отсутствие поиска;
6. лицензионная версия УДК (№ 4), не очень удобна в использовании из-за отсутствия навигации по индексам;
7. наиболее удобной для работы как систематизатора, так и читателя, являются базы данных, выпускаемые ГПНТБ России и Ассоциацией ЭБНИТ (№ 7) совместно с РГБ и ВИНТИ РАН, поскольку являются полностью лицензионными, русифицированными, имеют расширенные возможности поиска по индексам и ключевым словам, в том числе поиск по исключенным индексам для БД УДК, и удобную навигацию.

Список литературы

1. УДК – Универсальная десятичная классификация. URL: http://www2.viniti.ru/index.php?option=com_content&view=article&id=41:udc (Дата обращения: 04.02.2016)
2. UDC Summary. URL: <http://www.udcsummary.info/php/index.php> (Дата обращения: 04.02.2016)
3. UDC Online. URL: <http://www.udc-hub.com/en/login.php> (Дата обращения: 04.02.2016)
4. Продукты семейства ИРБИС. URL: http://elnit.org/index.php?option=com_content&view=article&id=35&Itemid=452 (Дата обращения: 04.02.2016)