

**МЕТОДЫ ИНДЕКСИРОВАНИЯ В ЦНБ УРО РАН ДЛЯ
ПОЛНОТЫ ОТРАЖЕНИЯ НАУЧНОГО
ДОКУМЕНТОПОТОКА**

Зырянова Ольга Николаевна,

главный библиотекарь,

Центральная научная библиотека Уральского отделения

Российской академии наук

В статье описана технология индексирования документов, применяемая в ЦНБ УрО РАН. Рассмотрены основные понятия в области индексирования, вопросы применения методов индексирования для полноты отражения научного документопотока и оптимизации информационно-библиотечного обслуживания ученых УрО РАН.

Ключевые слова: индексирование, систематизация, предметизация, посткоординация, информационно-поисковые языки, ЦНБ УрО РАН

**INDEXING METHODS IN THE CENTRAL SCIENTIFIC
LIBRARY OF THE URAL BRANCH OF THE RUSSIAN
ACADEMY OF SCIENCES FOR COMPLETE REFLECTION
OF THE SCIENTIFIC DOCUMENT FLOW**

*Ziryanova Olga Nikolaevna, head librarian,
Central Scientific Library of the Ural Branch
of the Russian Academy of Sciences*

The article describes the technology of indexing documents used in the Central Scientific Library of the Ural Branch of the Russian Academy of Sciences. It considers the basic concepts in the field of indexing, using indexing methods to fully reflect the scientific document flow and optimize information and library services for scientists of the Ural Branch of the Russian Academy of Sciences.

Keywords: indexing, systematization, objectification, post-coordination, information retrieval languages, Central Scientific Library of the Ural Branch of the Russian Academy of Sciences

Курс на стремительное развитие науки и сферы научной информации, обозначенный в Стратегии научно-технологического развития Российской Федерации [1], постоянное увеличение информационных ресурсов во всех отраслях человеческой деятельности и, соответственно, появление новых терминов требует их структурирования и систематизации для оперативного и эффективного поиска необходимой информации. Индексирование предназначено для создания поискового образа документа (ПОД) и ориентации пользователя в растущем объеме научной информации. Основной задачей применения классификационных и предметизационных систем в научных библиотеках является создание справочно-поискового аппарата, служащего инструментом доступа к документным фондам. Применение систем классификации и рубрикаторов позволяет отслеживать изменения в тематических областях науки, устанавливать соответствие фонда библиотеки профильному документопотоку и потребностям пользователей. Операции индексирования упорядочивают процессы комплектования, организации и хранения фонда, помогают при отборе документов, статистическом учете и анализе обращаемости фонда [2, с. 1; 3].

Индексирование представляет собой массовый технологический процесс, в ходе которого приводятся в действие все реализуемые в библиотеке лингвистические средства. В зависимости от типа используемого информационно-поискового языка оно может трансформироваться в такие процессы, как систематизация, предметизация, координатное индексирование. Информационно-поисковый язык (ИПЯ) – формализованный искусственный язык, предназначенный для индексирования документов, информационных запросов и описания фактов с целью последующего хранения и поиска [4]. В любом случае индексирование рассматривается как важнейший способ раскрытия содержания документов, помогает в их поиске.

Процесс индексирования регламентируется следующими нормативными документами:

- ГОСТ 7.74-96 «Информационно-поисковые языки. Термины и определения» [5];
- ГОСТ 7.59-2003 «СИБИД. Индексирование документов. Общие требования к систематизации и предметизации» [6];
- ГОСТ Р 7.0.66-2010 «Система стандартов по информации, библиотечному и издательскому делу. Индексирование документов. Общие требования к координатному индексированию» [7].

В Центральной научной библиотеке Уральского отделения Российской академии наук (ЦНБ УрО РАН) применяются методики

предкоординатного и координатного индексирования документов (систематизация, предметизация, индексирование ключевыми словами). Процесс индексирования включает следующие этапы:

- анализ содержания документа, определение его целевого назначения и формы;
- подготовка предварительного индексационного решения;
- перевод принятого решения на следующие информационно-поисковые языки: Библиотечно-библиографическая классификация (ББК), Государственный рубрикатор научно-технической информации (ГРНТИ) и принятые в ЦНБ УрО РАН таблицы библиотечно-библиографической классификации (автор Сухманева Е. Г.);
- составление предметных рубрик Алфавитно-предметного указателя (АПУ) и Центрального систематического каталога (ЦСК).

Научную обработку в ЦНБ УрО РАН проходят различные виды документов, поступающих в библиотечные фонды: произведения печати (книги и брошюры), продолжающиеся издания, периодические издания, диссертации, авторефераты диссертаций, препринты, издания на электронных носителях. Объектом индексирования могут быть:

- документ, в качестве самостоятельной единицы поступающий в библиотечный фонд;
- часть документа (например, статья из журнала);
- несколько отдельных документов (например, многотомное издание).

Предкоординатным называется индексирование, осуществляющееся на базе ИПЯ предкоординатного типа, преимущественно классификационных ИПЯ. Его отличительная особенность заключается в том, что лексические единицы индексируемого сообщения образуют линейную запись. Порядок их следования является жестким и устанавливается индексатором в ходе индексирования документов до поступления запроса («предкоординация», т. е. предварительность индексирования). Это обуславливает использование элементов позиционной грамматики, обеспечивающей строгое чередование элементов поискового образа в соответствии с классификационной формулой [4]. Классификационная формула (КФ) – «унифицированный порядок расположения классификационных индексов, определяющий классификационное решение» [6].

В рамках предкоординатного индексирования различают две его разновидности: систематизацию и предметизацию, основанные на двух основных принципах индексирования – классификационном и предметизационном. Систематизация – это «присвоение данным или документам классификационных индексов в соответствии с

правилами какого-либо классификационного ИПЯ» [4]. Систематизация осуществляется по таблицам классификации, отражающим специфические условия и требования научной библиотеки и информационные потребности пользователей.

Основным элементом таблиц классификации является классификационное деление (запись класса), в состав которого входят:

- классификационный индекс;
- словесная формулировка (описание класса);
- методические указания;
- справочно-ссылочный аппарат.

Классификационный индекс состоит из одной или нескольких лексических единиц (ЛЕ), которые следует выражать цифровыми, буквенными или смешанными (буквенно-цифровыми) обозначениями с использованием синтаксических знаков. В зависимости от количества ЛЕ в классификационном индексе различают простой, сложный или составной классификационные индексы. Лексическая единица (ЛЕ) – обозначение отдельного понятия, принятое в информационно-поисковом языке и неделимое в этой функции. Лексические единицы могут представлять собой принятые в естественном языке слова, устойчивые словосочетания, аббревиатуры, символы, даты, общепринятые сокращения, лексически значимые компоненты сложных слов [6].

Предварительный разбор поступивших на систематизацию документов производится по языковому признаку. Анализ содержания документа включает:

- определение предмета (одного или нескольких), рассматриваемого в документе;
- определение степени равнозначности предметов, рассматриваемых в документе, при наличии нескольких предметов в документе;
- определение основного и вспомогательных предметов, при наличии нескольких неравнозначных предметов, рассматриваемых в документе;
- определение взаимовлияния одного предмета на другой;
- определение степени актуальности каждого из предметов документа;
- определение аспектов исследования предметов документа;
- определение целевого и читательского назначения документа.

По результатам анализа принимается одно из классификационных решений:

- определение (подбор) по таблицам классификации (с привлечением АПУ) точного и детального табличного (готового) индек-

са, характеризующего документ во всех аспектах рассмотрения его предмета;

– составление сложного или составного индекса или внесение редакционного уточнения в раздел таблиц – в пределах действующих рубрик, в случае отсутствия готового индекса в таблицах классификации.

В ЦНБ УрО РАН применяется Библиотечно-библиографическая классификация (ББК). Научно-библиографический отдел Центральной научной библиотеки УрО РАН при выполнении запросов читателей использует также Межгосударственный классификатор стандартов (МГК), Международную патентную классификацию (МПК).

Сегодня в ЦНБ УрО РАН используется шесть видов таблиц ББК:

1. Полный вариант ББК для научных библиотек в 30 т. (1963–1992 гг.);
2. Сокращенный вариант ББК для научных библиотек в 5 выпусках (1970–1972 гг.);
3. ББК для областных библиотек в 4 т. (1980–1983 гг.);
4. ББК. Рабочие таблицы для массовых библиотек в 1 т. (1997 г.) (Культура. Культурология; Этика);
5. Таблицы ЦНБ УрО РАН 1963 г. (автор Сухманева Е. Г.): Естественные науки (Физико-математические науки; Химические науки; Науки о Земле);
6. ББК. Средние таблицы (Вып. 1–7, 2001–2018).

Издания с 2001 г. систематизируются по Средним таблицам ББК с буквенным обозначением делений. Внутри каждого из основных отделов классификационной схемы образуются подотделы, внутри каждого из подотделов – подразделения для отдельных частных вопросов, относящихся к данной отрасли знания. В результате создается стройная система делений с последовательным переходом от общих вопросов к более частным.

Предметизация – это «предметное индексирование средствами языка предметных рубрик». Предметная рубрика – это краткое словесное выражение предмета (темы) документа (запроса), основных аспектов его исследования, формы и назначения. Предметная рубрика определяет основную тематику документов, собранных под этой рубрикой. Структура предметной рубрики включает заголовок и подзаголовок. Язык предметных рубрик обеспечивает отражение в поисковом образе документа предметов и аспектов их рассмотрения, предназначен для проведения предметного поиска информации [4].

Предметизационный принцип индексирования базируется на представлении содержания документа в понятийной системе опре-

деленного естественного языка и на использовании в качестве терминов индексирования лексических единиц естественного языка, что обеспечивает возможность организации информационного поиска по алфавитному признаку. Если систематизация осуществляется на основе таблиц классификации, рубрикаторов, классификаторов, то предметизация предполагает наличие словаря или списка предметных рубрик, из которого и выбираются необходимые для отражения документа предметные рубрики. Список основных предметных рубрик для отражения документов в фонде ЦНБ УрО РАН представлен в таблице 1.

Таблица 1
Основные предметные рубрики в ЦНБ УрО РАН

<i>№</i>	<i>Рубрика</i>	<i>Классификационные индексы до 2001 г.</i>	<i>Классификационные индексы с 2001 г.</i>
	АСТРОНОМИЯ	52	В6
	БИБЛИОТЕЧНОЕ ДЕЛО	02.3; 02.5	Ч83; Ч85
	БИОЛОГИЧЕСКИЕ НАУКИ	57	Е
	ВОЕННАЯ НАУКА. ВОЕННОЕ ДЕЛО	355	Ц
	ГОРНОЕ ДЕЛО	65	И
	ЕСТЕСТВЕННЫЕ НАУКИ	50; 502.7	Б
	ЗДРАВООХРАНЕНИЕ. МЕДИЦИНСКИЕ НАУКИ	61	Р
	ИСКУССТВО. ИСКУССТВОЗНАНИЕ	7	Щ
	ИСТОРИЯ. ИСТОРИЧЕСКИЕ НАУКИ	9, 9.1/9.5	Т, Т1/5
	КУЛЬТУРА	371; 377; 379	Ч1, Ч7, Ч9
	ЛИТЕРАТУРА УНИВЕРСАЛЬНОГО СОДЕРЖАНИЯ	016; 03	016; 03
	ЛИТЕРАТУРОВЕДЕНИЕ	43	III3
	МАТЕМАТИКА	51	В1
	НАУКА. НАУКОВЕДЕНИЕ	001.2	Ч2
	НАУКИ О ЗЕМЛЕ	548; 549; 55; 91	Д
	НАУЧНО-ИНФОРМАЦИОННАЯ ДЕЯТЕЛЬНОСТЬ	001.3	Ч86
	ОБРАЗОВАНИЕ. ПЕДАГОГИЧЕСКАЯ НАУКА	374	Ч4
	ОБЩЕСТВЕННЫЕ НАУКИ	3; 3.5; 3.6; 3.7	С, С5/9
	ПИЩЕВЫЕ ПРОИЗВОДСТВА		Л8/9
	ПОЛИТИКА. ПОЛИТИЧЕСКАЯ НАУКА	32	Ф
	ПРАВО. ЮРИДИЧЕСКИЕ НАУКИ	34	Х
	ПСИХОЛОГИЯ	1.5	Ю9

РАДИОЭЛЕКТРОНИКА	648–649	<u>38/9</u>
РЕЛИГИЯ. АТЕИЗМ	2	Э
СЕЛЬСКОЕ И ЛЕСНОЕ ХОЗЯЙСТВО	63	П
СРЕДСТВА МАССОВОЙ ИНФОРМАЦИИ. КНИЖНОЕ ДЕЛО	376	Ч6
СТРОИТЕЛЬСТВО	69	Н
ТЕХНИКА. ТЕХНИЧЕСКИЕ НАУКИ	62, 67	Ж; М1, М2, М8, М9
ТЕХНОЛОГИЯ МЕТАЛЛОВ. МАШИНОСТРОЕНИЕ. ПРИБОРОСТРОЕНИЕ	669	К
ТРАНСПОРТ	68	<u>О</u>
ФИЗИКА	53	В3
ФИЗИЧЕСКАЯ КУЛЬТУРА И СПОРТ	7А	Ч5
ФИЛОЛОГИЧЕСКИЕ НАУКИ	40	III0
ФИЛОСОФСКИЕ НАУКИ	1–1.3; 1.6–1.8	Ю0/8
ФОЛЬКЛОР. ФОЛЬКЛORИСТИКА	42	III2
ХИМИЧЕСКАЯ ТЕХНОЛОГИЯ	661–668	Л1/7
ХИМИЧЕСКИЕ НАУКИ	54	Г
ХУДОЖЕСТВЕННАЯ ЛИТЕРАТУРА	P1, P2	III4
ЭКОНОМИКА. ЭКОНОМИЧЕСКИЕ НАУКИ	33	У
ЭНЕРГЕТИКА	64–647	<u>31/7</u>
ЯЗЫКОЗНАНИЕ	41	III1

Посткоординатным, или координатным индексированием называется индексирование текста документа путем перечисления ключевых слов, или дескрипторов, логическое произведение которых выражает основной смысл содержания данного текста, образуя его поисковый образ. Разница между ключевым словом и дескриптором обусловлена их принадлежностью к разным типам информационно-поисковых языков [4].

Ключевое слово – это слово или словосочетание естественного языка, которое с точки зрения информационного поиска несет наибольшую смысловую нагрузку. Оно выделяется из текста документа или запроса. В качестве ключевых слов используются лексические единицы, представляющие слова (унитермы), терминологические словосочетания, аббревиатуры, численные характеристики, хронологические данные, имена собственные, символические обозначения. Лексические единицы поискового образа при таком индексировании задаются списком, перечнем и сочетаются друг с другом после поступления запроса («посткоординация», т. е. последующая координация).

В электронные каталоги и базы данных ЦНБ УрО РАН ключевые слова заводятся прописными буквами. На первом месте в ключевых словах следует существительное (например, АВАРИЯ ЧЕРНОБЫЛЬСКАЯ), за исключением общепринятых словосочетаний, где на первом месте – прилагательное (например, ВЕЛИКАЯ ОТЕЧЕСТВЕННАЯ ВОЙНА, БАНКОВСКОЕ ДЕЛО). Ключевые слова употребляются во множественном числе (например, БАНКИ РОССИИ), за исключением слов, употребляющихся только в единственном числе (например, РУССКАЯ НАРОДНАЯ МУЗЫКА).

В ключевых словах не допускается употребление знаков препинания (точек, запятых) и сокращений слов (например, межд. отнош.). Если ключевое слово включает фамилию автора – она выносится на первое место (например, АБЕЛЯ ТЕОРЕМА). Нельзя в качестве ключевого слова вводить целую фразу (тему реферата), так как фраза не будет являться ключевым словом и поиск может оказаться безрезультатным. Для уточнения понятия в круглых скобках может быть введена тема, географическое понятие и т. д. (например, БЕЛКИ (ХИМИЯ ОРГАНИЧЕСКАЯ)). При вводе ключевого слова в виде аббревиатуры желательно ввести его полное наименование (например, ЯМР; ЯДЕРНО-МАГНИТНЫЙ РЕЗОНАНС).

Грамотное индексирование необходимо при поиске информации и при расстановке библиотечных фондов. В целом, применяемая в ЦНБ УрО РАН современная технология индексирования способствует более эффективному использованию фондов библиотеки в научной деятельности и повышает качественный уровень информационно-библиотечного обслуживания пользователей [8, с. 93]. Развитие методов и систем индексирования позволяет более полно и объективно отражать изменяющийся научный документопоток, способствует правильной идентификации и релевантному поиску документов, поэтому является крайне значимым в процессе построения и организации единого научного информационного пространства.

Список источников

1. О Стратегии научно-технологического развития Российской Федерации : указ Президента РФ от 01.12.2016. – URL : <http://www.kremlin.ru/acts/bank/41449> (дата обращения: 12.12.2021).
2. Прозоров И. Е. Индексы ББК в комплексе лингвистических средств качественной каталогизационной записи // Румянцевские чтения – 2017: 500-летие издания первой славянской Библии Франциска Скорины: становление и развитие культуры книгопечатания : мат-лы междунар. науч.-практ. конф. – Москва, 2017. – С. 51–65.

3. Урбанская Г. Г. Отображение терминологии по экономике АПК в тезаурусе ЦСХНБ // Культура: теория и практика. – 2019. – № 6. – С. 1–10.
4. Гендина Н. И. Лингвистические средства библиотечно-информационных технологий : учебник для вузов. – Санкт-Петербург : Профессия, 2015. – 439 с.
5. ГОСТ 7.74–96 Информационно-поисковые языки. Термины и определения // Стандарты по библиотечно-информационной деятельности // Стандарты по библиотечно-информационной деятельности. – Санкт-Петербург, 2003. – С. 318–343.
6. ГОСТ 7.59–2003 Индексирование документов. Общие требования к систематизации и предметизации // Сборник основных российских стандартов по библиотечно-информационной деятельности. – Санкт-Петербург, 2006. – С. 258–265.
7. ГОСТ Р 7.0.66-2010 Система стандартов по информации, библиотечному и издательскому делу. Индексирование документов. Общие требования к координатному индексированию. – URL : <https://internet-law.ru/gosts/gost/50758/> (дата обращения: 12.12.2021)
8. Зырянова О. Н. Индексирование документов в ЦНБ УрО РАН // Информационное обеспечение науки: новые технологии. – Екатеринбург, 2018. – С. 87–94.

References

1. *O Strategii nauchno-tehnologicheskogo razvitiya Rossijskoj Federacii : ukaz Prezidenta RF ot 01.12.2016.* [On the Strategy of Scientific and Technological development of the Russian Federation : Decree of the President of the Russian Federation dated 01.12.2016]. URL : <http://www.kremlin.ru/acts/bank/41449> (accessed: 12.12.2021). (In Russ.).
2. Prozorov I. E. *Indeksy BBK v komplekse lingvisticheskikh sredstv kachestvennoj katalogizacionnoj zapisi* [Classification numbers of the Library-Bibliographical Classification in the complex of linguistic means for quality catalogue recording]. *Rumyantsevskie chteniya – 2017: 500-letie knigopechataniya : materiali mezhdunarodnoj nauchno-prakticheskoy konferentsii.* Moscow, 2017, pp. 51–65. (In Russ.).
3. Urbanskaya G. G. *Otobrazhenie terminologii po ekonomike APK v tezauruse CSKHNB* [Displaying terminology on the AIC economy in the CSALthesaurus]. *Kultura: teoriya i praktika*, 2019, no. 6, pp. 1–10. (In Russ.).
4. Gendina N. I. *Lingvisticheskie sredstva bibliotechno-informacionnyh tekhnologij : uchebnik dlya vuzov* [Linguistic means of library and information technologies : textbook for universities]. St. Petersburg, *Professiya*, 2015, 439 p. (In Russ.).

5. GOST 7.74–96 *Informacionno-poiskovye yazyki. Terminy i opredeleniya* [Information retrieval languages. Terms and definitions : GOST 7.74-96]. *Standarty po bibliotechno-informacionnoj deyatel'nosti*. St. Petersburg, 2003, pp. 318–343. (In Russ).
6. GOST 7.59–2003 *Indeksirovanie dokumentov. Obshchie trebovaniya k sistematizacii i predmetizacii* [Indexing of documents. General requirements for systematization and objectification : GOST 7.59-2003]. *Sbornik osnovnykh rossijskikh standartov po bibliotechno-informacionnoj deyatel'nosti*. St. Petersburg, 2006, pp. 258–265. (In Russ).
7. GOSTR 7.0.66-2010 *Sistema standartov po informacii, biblioteknomu i izdatel'skomu delu. Indeksirovanie dokumentov. Obshchie trebovaniya k koordinatnomu indeksirovaniyu* [System of standards for information, library and publishing. Indexing of documents. General requirements for coordinate indexing : GOST R 7.0.66-2010]. URL : <https://internet-law.ru/gosts/gost/50758/> (accessed: 12.12.2021). (In Russ).
8. Ziryanova O. N. *Indeksirovanie dokumentov v CNB UrO RAN* [Indexing of documents in the Central scientific library of the Ural branch of the Russian academy of sciences]. *Informatsionnoe obespechenie nauki: novye tehnologii*, Yekaterinburg, 2018, pp. 87–94. (In Russ).