УДК 025.4:004
ББК 78.653.4
DOI 10.25281/0869-608X-2020-69-3-243-260

О.А. ЛАВРЁНОВА, А.А. ВИНБЕРГ

Современные пользователи библиотек и пространство связанных открытых данных

Реферат. Задача любой библиотеки — обеспечить высокое качество и общедоступность средств поиска информации. В статье описывается реализованный Российской государственной библиотекой (РГБ) проект представления Библиотечно-библиографической классификации (ББК) в виде сетевой системы организации знаний. Цель проекта — поддержка контента и предоставление средств обеспечения интероперабельности с другими ресурсами той же природы, т. е. со словарями связанных данных (Linked Data vocabularies) в среде Глобальной сети. Проект частично поддержан Российским фондом фундаментальных исследований. В качестве основного ресурса данных для Классификационной системы организации знаний был выбран Генеральный систематический каталог РГБ. Семантика каждого классификационного индекса выражается полной цепочкой его словесных формулировок (расшифровок), а не только формулировкой последнего уровня индекса. Данные, преобразованные в схему описания ресурса (Resource Description Framework, RDF) на основе модели Простой системы организации знаний (Simple Knowledge Organization System, SKOS), были загружены в семантическое хранилище для соответствующей обработки с использованием языка запросов SPARQL. РГБ опубликовала Классификационную систему (https://lod.rsl.ru) в среде связанных открытых данных (Linked Open Data, LOD) для обогащения поиска в электронном каталоге. В настоящее время ведется работа в направлении реализации взаимодействия с другими словарями LOD. Для дифференциации типов связей элементов (концептов) различных словарей (например, ББК с УДК, MeSH, нормативными файлами), представленных на базе модели SKOS, используются специальные метки. Концептуальные структуры ведущих классификаций имеют принципиальные различия. Поэтому формирование соответствий между концептами представляется реальным только на основе лексико-синтаксического анализа словесных формулировок индексов и требует серьезных исследований.

Ольга Александровна Лаврёнова,
Российская государственная библиотека,
отдел методического и лингвистического обеспечения информационных систем,
начальник
Воздвиженка ул., д. 3/5,
Москва, 119019, Россия
кандидат филологических наук,
старший научный сотрудник
ORCID 0000-0002-4779-9222
E-mail: LavrenovaOA@rsl.ru

Андрей Анатольевич Винберг,
Российская государственная библиотека,
главный технолог
Воздвиженка ул., д. 3/5,
Москва, 119019, Россия
кандидат технических наук
E-mail: VinbergAA@rsl.ru

Авторы рассчитывают на сотрудничество с другими библиотеками, в первую очередь российскими, в плане создания общего семантического пространства связанных открытых словарей.

Ключевые слова: открытые связанные данные, Linked Data vocabularies, сетевые системы организации знаний, NKOS, Semantic Web, Семантическая паутина, Библиотечно-библиографическая классификация, информационный поиск, семантические дорожные карты, поисковые намерения, SKOS, RDF, виртуальный систематический каталог, обогащение запроса, электронные библиотеки, организация пространства библиотек.

Для цитирования: *Лаврёнова О.А., Винберг А.А.* Современные пользователи библиотек и пространство связанных открытых данных // Библиотековедение. 2020. Т. 69, № 3. С. 243—260. DOI: 10.25281/0869-608X-2020-69-3-243-260.

Кофе, комфорт, семантика и равноправие пользователей

Кофе, видимо, давно задумал завоевать пространство библиотек. В последние годы этот напиток ведет массовое наступление, ставя наличие в библиотеках кофейных заведений во главу угла, т. е. требуя себе собственный уютный уголок. Без хорошего кофе она уже не признается современной и в целом достойной будущего. Однако при всей любви авторов статьи к кофе, рассмотрим и другие требования к образцовым современным библиотекам.

Библиотеки должны предоставлять читателям комфортные условия для работы или отдыха в формах, соответствующих типу и возрасту их посетителей, а также требованиям времени (включая кофе и чай с плюшками). И это не шутка. Подчеркнем, что возраст пользователя — ключевой момент в данном вопросе, и особое внимание целесообразно обратить на условия для детей и молодежи, устраивая им в библиотеке «второй дом», где удобно, уютно, надежно, интересно, перспективно. Подчеркивание роли современной комфортной среды в библиотеках нередко вызывает ироничные комментарии, но фактически удобная мебель, адекватное обслуживание (и отличный кофе!) демонстрируют особое отношение к посетителю. Человек всегда предпочтет прийти туда, где ему оказывают больше уважения и внимания.

На такой же высокий уровень обслуживания пользователь вправе рассчитывать и в плане обеспечения информацией, которую он ищет в библиотеках наряду с удовлетворением духовных потребностей в целом. При этом информация, собираемая, хранимая и организуемая библиотеками на различных носителях, должна предоставляться посетителям и внешним пользователям на равных условиях вне зависимости от их местонахождения и жизненных обстоятельств. Это называется «информационным равенством граждан», достижение которого — пока дело будущего.

До сих пор поисковые возможности пользователей ограничены не только недостаточной доступностью Глобальной сети, но и технологиями, которые поддерживаются конкретными автоматизированными информационно-библиотечными системами. Особенно это влияет на семантический (тематический, предметный) поиск информационных ресурсов, который требует применения достаточно сложных лингвистических средств.

Библиотеки любого уровня всегда неуклонно стремились к обеспечению равенства читателей, используя для организации информации и фонда (а также, разумеется, для поиска) широко распространенные в конкретной стране или в мире библиотечно-библиографические классификации, системы предметизации, тезаурусы. Однако поиск на их основе с помощью компьютерных технологий нельзя назвать столь же распространенным. В результате в одних библиотеках имеется возможность полноценного семантического поиска с обогащением запросов более узкими по значению или ассоциативными понятиями по отношению к заданным в запросе лексическим единицам, их синонимами, различными грамматическими формами, а в других —приходится довольствоваться ключевыми словами или простыми предметными рубриками. Соответственно, требуется сделать сильные семантические средства поиска широкодоступными. Кроме того, они не могут быть одинаковыми для пользователей с разными типами информационных потребностей, которые зависят от рода деятельности людей, намерений и поставленных задач.

Цель проекта, которому посвящена настоящая статья, — поддержка контента и предостав-

ление средств обеспечения интероперабельности Классификационной системы организации знаний на основе Библиотечно-библиографической классификации (ББК) с другими словарями связанных данных (Linked Data vocabularies) в среде Семантической паутины (Semantic Web).

Пользователь пользователю рознь

Превалируют три разновидности поисковых намерений пользователей библиотек:

- 1-й тип — получение публикаций по заданному системе слову или сочетанию слов, требуется точность результата, полнота не принимается во внимание;
- 2-й тип — обнаружение системой максимума публикаций по заданной теме, важна полнота;
- 3-й тип — предварительный поиск в структурированной информации по интересующей области знания, построение маршрута поиска, например в классификаторе, программой и непосредственно пользователем, применение связей между науками.

В публичных выступлениях отечественных разработчиков электронных библиотек (ЭБ) просматривается тенденция отдавать предпочтение развитию полнотекстового поиска и его противопоставление якобы устаревшим классификациям, применяемым в библиотеках. Часто речь не идет даже о тезаурусах, которые нередко признаются простыми онтологиями, не говоря уже о классических онтологиях.

Действительно, пользователей, имеющих поисковые намерения 1-го типа и привыкших получать ответ в результате «одного клика», полнотекстовый поиск в системах Google и Yandex вполне устраивает. Однако намерения 2-го и 3-го типов требуют учета при поиске смысловых связей между словами запроса и лексическими единицами, отражающими смысловое содержание текста. Это условие обеспечивают библиотечно-библиографические классификации и словари типа тезаурусов. Реализация намерений 3-го типа предполагает наличие навигации в рамках модели представления (организации) знаний (классификации, тезауруса, онтологии) с использованием семантических связей. Поисковые системы (электронные каталоги (ЭК), ЭБ), реализующие классификации и тезаурусы, предоставляют пользователям так называемые «семантические дорожные карты» (Semantic Road Maps) [1]. Они позволяют построить маршруты от темы запроса к темам документов (электронных ресурсов), дополняя (обогащая) запросы другими поисковыми словами. Используя классификационные системы и тезаурусы, библиотеки ориентируются на все перечисленные типы намерений, предлагая пользователям соответствующие сервисы.

Формирование такого рода возможностей в рамках развития систем полнотекстового поиска на основе вероятностно-статистических методов и тем более создание так называемых систем искусственного интеллекта требуют кропотливого «обучения» программных продуктов на основе внушительного корпуса обработанных текстов. Перспективы их практического использования до сих пор нельзя назвать достаточно определенными.

Однако чашечка кофе прояснила голову удобно расположившегося в библиотеке читателя, он внезапно вспомнил, что изначально библиотека предназначена для получения информации, и обратился к поиску, рассчитывая на комфорт и в данном аспекте.

Полнотекстовый поиск и системы организации знаний

«В Интернете можно найти все» — трудно вспомнить более наивную фразу, чем эта. Во-первых, она исходит из ложной посылки «в Интернете есть все», что заведомо невозможно. Во-вторых, можно найти, а можно и не найти то, что требуется конкретному пользователю. Данная проблема касается преимущественно смыслового поиска (поиска по теме, предмету). Веб-технологии, в принципе, не предусматривали семантические связи между ресурсами [2]. Ссылки от одной веб-страницы к другой носят формальный характер, смысловой тип связи не обозначается, и поэтому тематический поиск весьма условен и случаен. Многие современные проекты, поисковые технологии которых ориентированы на формальные (вероятностно-статистические) методы поиска в электронных ресурсах, в частности в ЭБ, практически копируют и развивают технологию Всемирной паутины (например, на базе поисковых систем Google и Yandex).

Исторически библиотеки отличаются высокой организацией информации, если определить ее как совокупность знаний, предназначенных и подготовленных людьми для передачи другим людям в различной форме. Теория и практика

информационного поиска тщательно разработана и описана в публикациях достаточно давно [3; 4]. С тех пор свойства естественных языков, которыми передаются знания, ничуть не изменились, и библиотечные классификации, тезаурусы остались действующими семантическими моделями организации знаний [5; 6] в мировой библиотечной практике [7]. Не имеется никаких оснований отказываться от этих лингвистических средств в пользу исключительно полнотекстового поиска, у которого свое предназначение. Следует добавить, что нарушение структуры организации знаний в библиотеке означает создание условий для уничтожения возможности когда-либо найти многие документы в ее фондах, и чем крупнее библиотека, тем выше такого рода риски.

Продемонстрируем суть проблемы на простом примере.

Пример 1. Допустим, аспирант занимается актуальной проблемой сохранения языков малочисленных коренных народов России и для подготовки диссертации, посвященной исследованию некоторого явления в *угорской группе языков*, проводит поиск в Интернете по соответствующему словосочетанию. Если в полном тексте публикации, которая, в принципе, соответствует теме, встречаются только более узкие по смыслу термины, но нет заданного в запросе словосочетания, ресурс не будет найден. Таким образом, могут быть не обнаружены тексты, в частности, по следующим языкам: *хантыйский (остяцкий)*, *мансийский (вогульский)*, публикации по *обско-угорским языкам* в целом. Для обеспечения полноты результата придется провести поиски еще и по всем нужным языкам отдельно. Иерархические структуры классификаций [8] и тезаурусов [9; 10] избавляют от этой утомительной процедуры, поскольку в них моделируются знания о смысловых связях между понятиями.

Рассмотрим пример вывода результатов поиска по словосочетанию «*угорская группа языков*» (рис. 1) в Классификационной системе организации знаний с использованием отечественной Библиотечно-библиографической классификации (ББК) в ее полном варианте для научных библиотек [11]. Справа на рис. 1 мы видим индекс *Ш166.512* и последовательность словесных формулировок, содержащую все уровни иерархии: «*Филологические науки. Художественная литература -- Языкознание -- Языки мира -- Финно-угорские языки -- Угорская группа языков -- Обско-угорские языки -- Хантыйский (остяцкий) язык*». Каждую такую последовательность можно рассматривать как заранее разработанный маршрут для пользователя от заданного запроса (здесь «*угорская группа языков*») до всех более узких по содержанию тем и соответствующих ресурсов. Поэтому в международной практике классификационные структуры и стали называть «семантическими дорожными картами». Слева расположен фрагмент иерархического дерева ББК и выделены уровни иерархии, приведенные в цепочке словесных формулировок индекса (рис. 1).

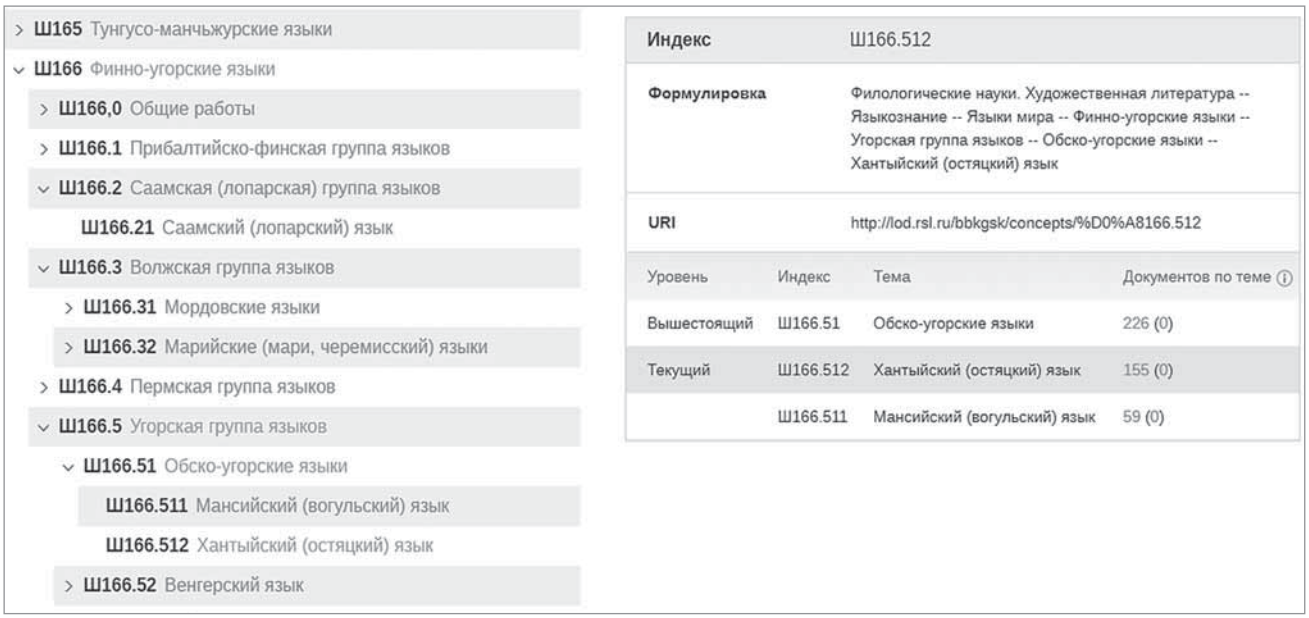


Рис. 1. Найденный по иерархической классификации индекс ББК для хантыйского (остяцкого) языка в результате поиска по словосочетанию «угорская группа языков» и перемещения вниз по иерархии делений (рубрик)

Отметим, что полнотекстовый поиск мог бы обнаружить все языки заданной в запросе группы в публикациях только при условии, что в текстах, кроме названий конкретных языков, будет упоминаться и название этой языковой группы. Вряд ли такое можно гарантировать априори.

Иерархические структуры классификаций и тезаурусов моделируют знания о смысловых связях между понятиями и обеспечивают программное добавление в запрос более узких по смыслу и ассоциативных лексических единиц, а также синонимов, существенно повышая полноту поиска. Такой процесс называют в библиотечной практике «обогащением» запроса (от англ. enrichment). В Российской государственной библиотеке (РГБ) автоматический учет смысловых отношений при поиске в ЭК и ЭБ обеспечен с самого начала их создания путем ввода непосредственно в библиографические записи словесных формулировок индексов [12; 13]. Связи обогащали запросы по умолчанию, без участия пользователей, существенно увеличивая полноту поиска. Затем был разработан пилотный проект навигатора по виртуальным разделителям Генерального систематического каталога (ГСК) РГБ на основе оцифровки, тестирования, редактирования, дополнения данных со всех 130 тыс. бумажных разделителей. В конечном счете РГБ реализовала проект Классификационной системы организации знаний на основе ББК [14].

Существует дополнительная проблема: в настоящее время можно считать достаточно перспективным поиск в организованных информационных ресурсах, например в ЭК, ЭБ, только из одной точки и едиными поисковыми средствами. Мало кто готов обращаться в целый ряд поисковых систем по одному и тому же запросу. Следовательно, требуется создавать соответствующие сервисы в библиотеках. Понятно, что поиск с одного компьютерного рабочего места по различным поисковым ресурсам реализуем при условии, что они связаны между собой, причем не просто ссылками, как в веб-среде, а смысловыми отношениями. Иными словами, каждая связь ресурсов должна иметь определенную семантику, причем в рамках глобальной сети.

Если вернуться к вопросу об информационном равенстве граждан, то уместно отметить, что специалисты РГБ решили не ограничиваться рамками собственной библиотеки и разработать технологии предоставления такого рода поиска другим библиотекам, причем не только использующим полные таблицы ББК, но и их средний вариант, а также другие классификации, предметные рубрики и тезаурусы. Основная идея заключается в объединении технологий тематического поиска в различных ЭК и ЭБ на базе созданной Классификационной системы организации знаний.

Для поддержки поиска в различных библиотеках страны РГБ опубликовала Классификационную систему организации знаний в среде связанных открытых данных (Linked Open Data, LOD) на сайте https://lod.rsl.ru под открытой лицензией РГБ, основанной на рекомендациях Creative Commons [15]. Научные исследования и разработки были осуществлены в 2015—2017 гг. при поддержке Российского фонда фундаментальных исследований (проект № 15-07-05265). В настоящее время Система функционально связана с ЭК РГБ.

При поиске библиографических записей по Классификационной системе рекомендуется учитывать следующие обстоятельства:

- на данный момент в модели представлены индексы с традиционных разделителей ГСК РГБ, имевшихся во время их оцифровки, и ряд добавленных для организации правильной структуры деревьев; планируется развитие иерархии вниз на основе данных в ЭК РГБ;
- ретроконверсия каталогов РГБ не завершена, поэтому при поиске по выбранному индексу пока невозможно гарантировать полноту выдачи библиографических записей, созданных до 2000 г., однако в таких случаях индекс полезно применить для дополнительного поиска в карточном ГСК РГБ;
- для поиска библиографических записей, подготовленных до 1961 г. (до внедрения таблиц ББК), предполагается восстановить электронный имидж-каталог на новой программной платформе.

Выход в свет, или Библиотеки в Semantic Web

Специальный уровень интернет-технологий Semantic Web (Семантическая паутина) [16] создан «над» технологиями WWW именно для семантического поиска в сетевом пространстве. Публикация систем организации знаний в виде словарей связанных данных в Semantic Web представляет собой реальную основу LOD-проектов [17; 18]. Принципиальное отличие технологии Semantic Web от технологии WWW состоит в том, что связи между ресурсами в первой имеют семантику. Например, зада-

ются связи «имеет вышестоящее (выше — ниже по иерархии)», «имеет ассоциативное понятие», «имеет название (заглавие, имя)», «имеет место работы», «является автором произведения» и т. д. В то же время, как отмечалось ранее, в веб-среде связи работают как простые ссылки без заданного заранее значения. В Сети существуют и создаются для отдельных задач пространства имен (Name Spaces) [19; 20], в которых фиксируются, в частности, виды связей между ресурсами. Разработчики конкретных LOD-словарей могут использовать элементы имеющихся в сети пространств имен или создать свое пространство по общим правилам данной технологии. При этом ресурсами считаются и целые документы, и отдельные лексические единицы, и предметные рубрики или классификационные индексы и т. д.

В современном мире библиотеки выводят читателя в открытую среду Semantic Web путем публикации в этой среде различных видов библиографических данных [21; 22]. Классификации, тезаурусы, системы предметных рубрик, нормативные (авторитетные) файлы рассматриваются в открытом сетевом пространстве как LOD-словари [23; 24]. Связи между ними позволяют обогащать (дополнять) запросы пользователей словами из других словарей и программно переходить к поиску ресурсов в системах других библиотек.

Формирование данных для LOD-технологий упрощенно выглядит следующим образом [25; 26]:

- каждый ресурс программно получает унифицированный идентификатор ресурса (Uniform Resource Identifier, URI) [27];
- выбираются и/или создаются пространства имен (абстрактные хранилища, множества);
- строятся утверждения о ресурсе в форме триплетов (троек) «субъект — предикат — объект» («ресурс — свойство ресурса — значение свойства») на базе схемы описания ресурса (Resource Description Framework, RDF) [28; 29].

Присвоение URI каждому индексу классификации — стандартный программный процесс. Формирование семантических связей на основе хорошо структурированных библиотеками файлов классификаций, как и любого нормативного/авторитетного файла, в том числе тезауруса, тоже выполняется программно. Более подробно данная технология и полное описание структуры элементов данных рассмотрены в публикациях [30; 31]. Для формирования LOD в форме триплетов «субъект — предикат — объект» было выбрано пространство имен Простой системы организации знаний (Simple Knowledge Organization System, SKOS) [32], специально предназначенное для представления таких словарей, как классификации, тезаурусы, онтологии. Затем была уточнена семантика каждого типа связей в рамках проекта.

В частности, концепт, имеющий индекс *Ш166.512* (см. пример 1 и рис. 1), получил URI: http://lod.rsl.ru/bbkgsk/concepts/%D0%A8166.512, а его вышестоящий индекс *Ш166.51* — URI: http://lod.rsl.ru/bbkgsk/concepts/%D0%A8166.51.

Можно сказать, что представление индекса в RDF — это описание соответствующего концепта как некоторой абстрактной сущности, некоего «смыслового содержания». Связь между URI концепта с его индексом обозначается как *skos:notation,* связь URI концепта с полной цепочкой словесных формулировок индекса — *skos:prefLabel*, а с URI вышестоящего концепта (фактически с URI вышестоящего индекса) — skos:broader и т. д. (см. пример 2).

В виде триплета последняя связь интерпретируется следующим образом: «*Ш166.512* (это субъект) — *имеет вышестоящим* (это предикат) — *Ш166.51* (это объект)».

Пример 2. Сокращенное представление концепта с индексом *Ш166.512* из примера 1 в RDF:

<skos:notation> Ш166.512 </skos:notation> (связь концепта как некоторого абстрактного понятия с его обозначением, т. е. индексом ББК)

<skos:prefLabel xml:lang=»ru»> *Филологические науки. Художественная литература -- Языкознание -- Языки мира -- Финно-угорские языки -- Угорская группа языков -- Обско-угорские языки -- Хантыйский (остяцкий) язык* </skos:prefLabel> (связь концепта с его названием на русском языке, т. е. с приведенной полной цепочкой словесных формулировок индекса)

<skos: broader xml:lang=»ru»> http://lod.rsl.ru/bbkgsk/concepts/%D0%A8166.51 </skos: broader> (связь описываемого концепта с URI вышестоящего концепта, практически с вышестоящим индексом *Ш166.51,* имеющим словесную формулировку «*Филологические науки. Художественная литература -- Языкознание -- Языки мира -- Финно-угорские языки -- Угорская группа языков -- Обско-угорские языки*»)

Аналогичным образом задаются ассоциативные связи «смотри также», связи с примеча-

ниями-примерами из методических указаний к индексам и с другими элементами ББК.

Представленные в RDF данные были загружены в семантическое хранилище для последующего манипулирования ими с помощью языка запросов SPARQL. Объем статьи не позволяет раскрыть назначение и структуру запросов данного типа [33; 34]. Важно подчеркнуть, что именно благодаря настройке структуры запроса SPARQL поиск любой сложности может осуществляться программно без участия человека как в одной библиотечной системе, так и в нескольких при условии публикации их словарей в среде LOD и установления связей между ресурсами.

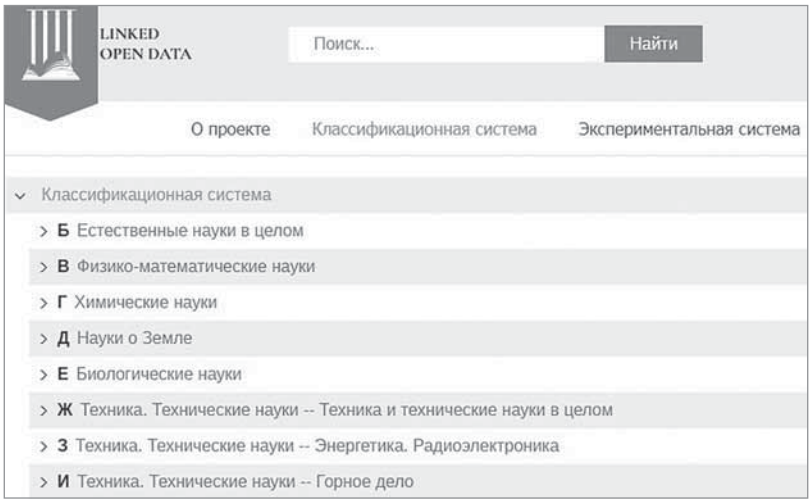


Рис. 2. Первая страница Классификационной системы организации знаний с начальным фрагментом иерархической структуры ББК

Пользователь может вести поиск в Классификационной системе [35] на основе ББК/ГСК [14] путем навигации по ее иерархической структуре, ассоциативным связям (рис. 2) или ввести произвольное сочетание слов в поисковую строку (рис. 3). По сути, эта технология поддерживает функции виртуального систематического каталога.

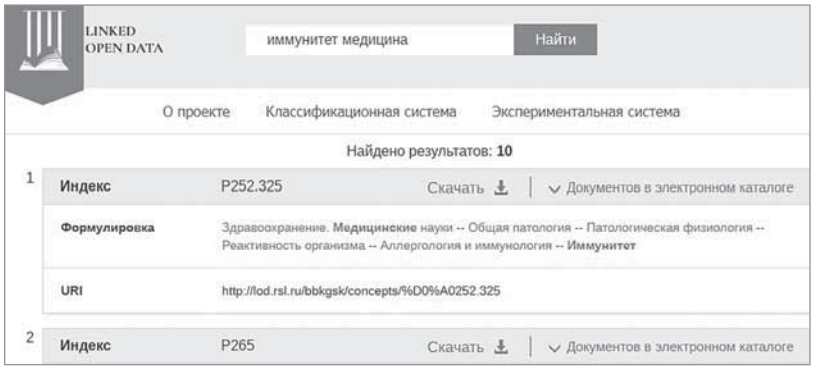


Рис. 3. Поиск в Классификационной системе организации знаний по теме «Иммунитет в медицине» (фрагмент выдачи результата)

Автоматическое обогащение запроса пользователя по технологии LOD [36; 37] на основе грамматических и семантических словарей происходит в Классификационной системе путем программного вовлечения в процесс поиска:

- эквивалентов лексических единиц из формулировок: форм словоизменения и словообразования, синонимов;
- более узких или ассоциативных индексов;
- слов из примечаний (указаний) к индексам в таблицах классификации;
- индексов других связанных классификаций или лексических единиц тезаурусов и предметных рубрик (естественно, после установления связей с соответствующими словарями).

Далее приведены примеры способов обогащения запросов пользователей.

Пример 3. При поиске по слову «*дети*» программно подключаются слова «*детский*», «*ребенок*», «*детство*», а также все падежные формы этих слов.

Пример 4. При поиске по индексу ББК запрос может быть автоматически дополнен индексами, расположенными ниже по иерархии или связанными ссылкой «смотри также».

В Классификационной системе построены иерархические связи от индекса *Щ* до индекса *Щ310.64 «Искусство. Искусствознание -- Музыка -- Теория музыки (теоретическое музыкознание) --Музыкальная акустика. Естественнонаучные основы музыки -- Акустика певческого голоса» (http://lod.rsl.ru/bbkgsk/concepts/%D0%A9310.64)*. Благодаря этому на запрос «*теория музыки*» система, в частности, найдет и публикации по теме «*Акустика певческого голоса*» с индексом *Щ310.64*.

Поскольку в системе установлена ассоциативная связь найденного индекса *Щ310.64* с индексом *Щ314.3-1 «Голосовой аппарат. Гигиена голоса»*, то на экране появится ссылка «*смотри также*», по которой доступен дополнительный поиск на основе индекса. *Щ314.3-1*.

Пример 5. Если системой найден индекс *Щ334.076.3 «Искусство. Искусствознание -- Театр -- Отдельные виды театра -- Драматический театр -- Теория дра-*

матического театра -- Творческие компоненты и выразительные средства драматического спектакля -- Искусство актера -- Сценическая речь» (http://lod.rsl.ru/bbkgsk/concepts/%D0%A9334.076.3), то системой будет предложен дополнительный поиск по словам из примеров в методических указаниях к индексу («*постановка голоса, дикция, интонация*» и все грамматические формы этих слов).

Поисковые процессы подробно описаны в разделе «О проекте» на сайте Классификационной системы [14], а также в статье [31].

В настоящее время в РГБ рассматриваются возможности и способы установления связей со следующими словарями:

- средний вариант таблиц ББК [38];
- таблицы УДК (UDC Summary) [39] (опубликовано в LOD в открытом доступе около 2,5 тыс. верхних рубрик; пример эксперимента РГБ по установлению соответствий индексов УДК с индексами ББК см.: https://lod.rsl.ru/experimental/bbkgsk/concepts/Щ32/);
- нормативный файл-тезаурус географических названий РГБ [40; 41];
- тезаурус по медицине MeSH [42] (опубликован в среде LOD).

Формирование связей между элементами различных словарей — непростая задача, так как в большинстве случаев прямые смысловые соответствия между ресурсами словарей построить сложно. Решать ее целесообразно на лексико-синтаксическом уровне. Причина состоит в том, что такие словари имеют отличные друг от друга концептуальные основы и используют принципиально разные способы организации знаний. В настоящее время небольшой коллектив в РГБ занимается данными разработками, но требуется государственная поддержка сложных и трудоемких работ, участие в них других библиотек.

Убедительным аргументом в пользу необходимости развития данного направления библиотечной деятельности в России можно считать серьезные масштабы зарубежных исследований и внедренных проектов по публикации LOD-словарей в Semantic Web за последнее десятилетие, особенно в национальных библиотеках. При этом их элементы (электронные ресурсы) связаны между собой и обеспечивают программный переход при поиске из каталогов одной библиотеки в каталоги других.

Для установления семантических связей между концептами из различных LOD-словарей используются их обозначения из пространства имен SKOS:

- *skos:exactMatch* — точное соответствие;
- *skos:closeMatch* — близкое по значению;
- *skos:relatedMatch* — имеет отношение;
- *skos:broadMatch* — соответствие в качестве вышестоящего;
- skos:narrowMatch — соответствие в качестве нижестоящего.

Разумеется, связи формируются между URI концептов.

Пример 6. В полном варианте ББК имеется 10 концептов с понятием «*иммунитет (в медицине)*» в различных аспектах (рис. 3). Один из концептов выражен индексом *Р252.325* и цепочкой его словесных формулировок:

ББК: *Р252.325 Здравоохранение. Медицинские науки -- Общая патология -- Патологическая физиология -- Реактивность организма -- Аллергология и иммунология -- Иммунитет.*

Для понятия «*иммунитет (в медицине)*» представлены страницы различных LOD-словарей: Wikidata (рис. 4), Библиотеки Конгресса США (рис. 5), Национальной библиотеки Франции (рис. 6), Национальной библиотеки Германии (рис. 7). Они обеспечивают возможность создания связей с лексическими

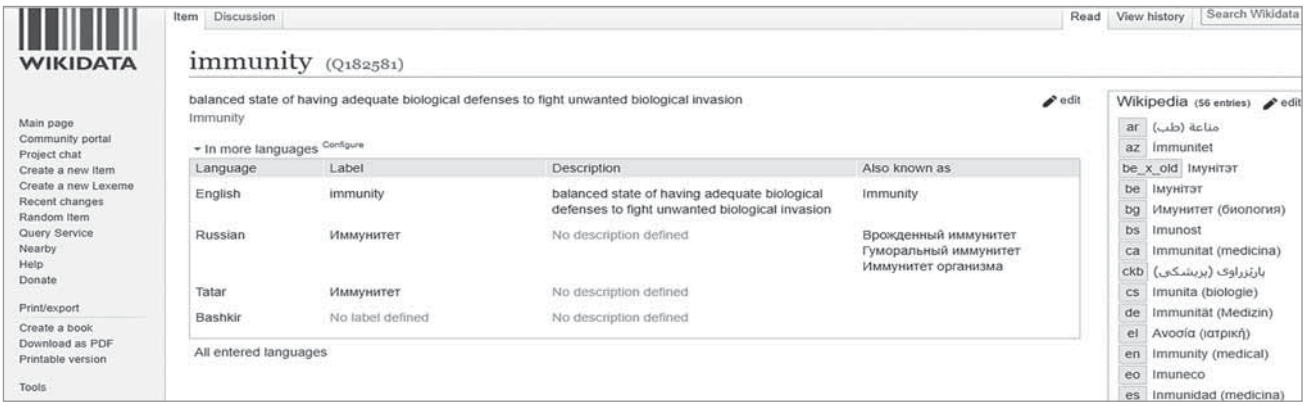
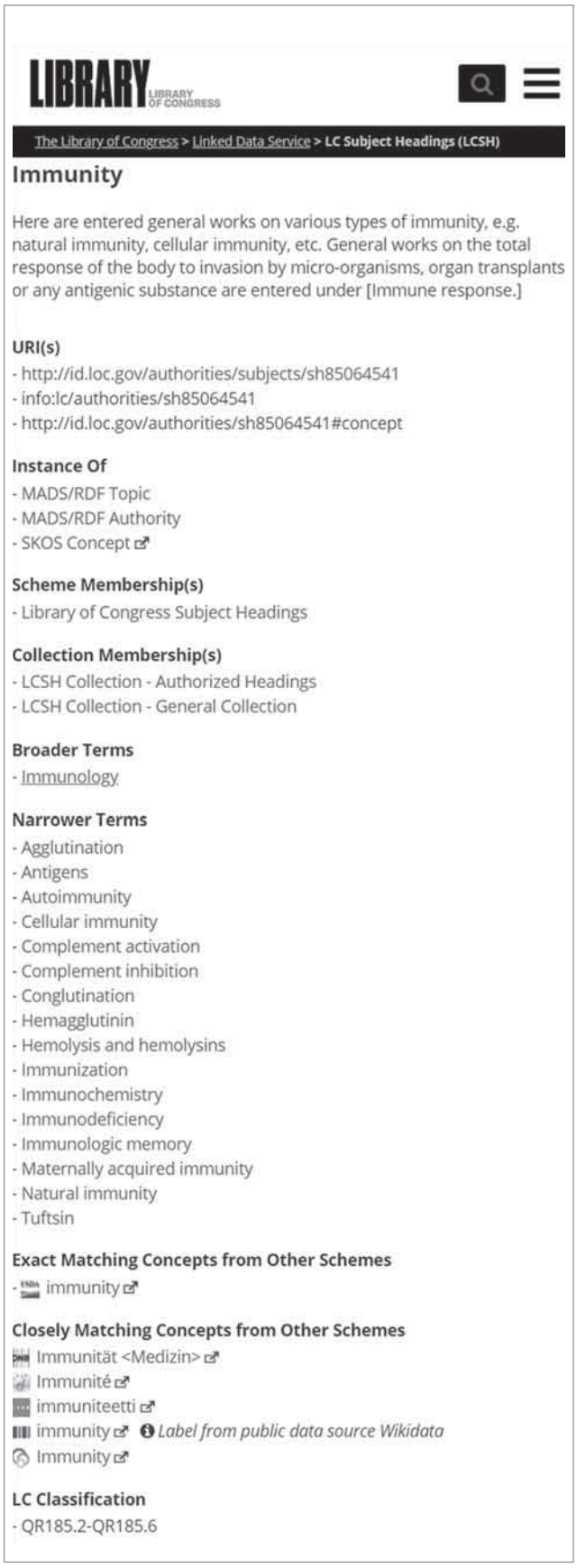


Рис. 4. «Иммунитет» в Wikidata (Википедия на английском языке в среде LOD)



LIBRARY LIBRARY OF CONGRESS

The Library of Congress > Linked Data Service > LC Subject Headings (LCSH)

Immunity

Here are entered general works on various types of immunity, e.g. natural immunity, cellular immunity, etc. General works on the total response of the body to invasion by micro-organisms, organ transplants or any antigenic substance are entered under [Immune response.]

URI(s)
- http://id.loc.gov/authorities/subjects/sh85064541
- info:lc/authorities/sh85064541
- http://id.loc.gov/authorities/sh85064541#concept

Instance Of
- MADS/RDF Topic
- MADS/RDF Authority
- SKOS Concept

Scheme Membership(s)
- Library of Congress Subject Headings

Collection Membership(s)
- LCSH Collection - Authorized Headings
- LCSH Collection - General Collection

Broader Terms
- Immunology

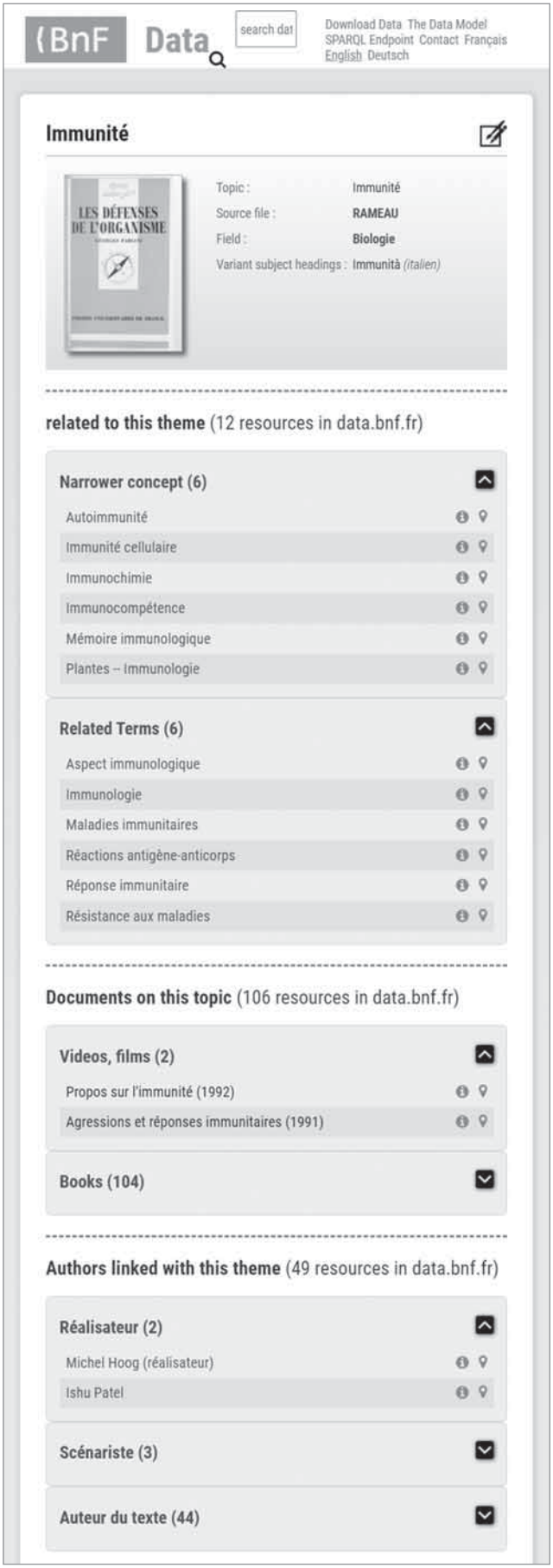
Narrower Terms
- Agglutination
- Antigens
- Autoimmunity
- Cellular immunity
- Complement activation
- Complement inhibition
- Conglutination
- Hemagglutinin
- Hemolysis and hemolysins
- Immunization
- Immunochemistry
- Immunodeficiency
- Immunologic memory
- Maternally acquired immunity
- Natural immunity
- Tuftsin

Exact Matching Concepts from Other Schemes
- immunity

Closely Matching Concepts from Other Schemes
- Immunität <Medizin>
- Immunité
- immuniteetti
- immunity *Label from public data source Wikidata*
- Immunity

LC Classification
- QR185.2-QR185.6

Рис. 5. «Иммунитет» в нормативном файле предметных рубрик Библиотеки Конгресса США. Указаны связи со словарями других библиотек



BnF Data search dat

Download Data The Data Model SPARQL Endpoint Contact Français English Deutsch

Immunité

Topic : Immunité
Source file : RAMEAU
Field : Biologie
Variant subject headings : Immunità *(italien)*

related to this theme (12 resources in data.bnf.fr)

Narrower concept (6)
- Autoimmunité
- Immunité cellulaire
- Immunochimie
- Immunocompétence
- Mémoire immunologique
- Plantes -- Immunologie

Related Terms (6)
- Aspect immunologique
- Immunologie
- Maladies immunitaires
- Réactions antigène-anticorps
- Réponse immunitaire
- Résistance aux maladies

Documents on this topic (106 resources in data.bnf.fr)

Videos, films (2)
- Propos sur l'immunité (1992)
- Agressions et réponses immunitaires (1991)

Books (104)

Authors linked with this theme (49 resources in data.bnf.fr)

Réalisateur (2)
- Michel Hoog (réalisateur)
- Ishu Patel

Scénariste (3)

Auteur du texte (44)

Рис. 6. «Иммунитет» — тематический поиск на сайте Национальной библиотеки Франции

Link zu diesem Datensatz	http://d-nb.info/gnd/4120476-1
Sachbegriff	Immunität <Medizin>
Quelle	M 3., SWL
Erläuterungen	Definition: Immunität gegen Antigene
Oberbegriffe	Infektionsabwehr
Thematischer Bezug	Verwandter Begriff: Resistenz
DDC-Notation	571.96 616.079 636.0896079
Systematik	27.8b Immunologie
Typ	Allgemeinbegriff (saz)
Andere Normdaten	LCSH: Immunity RAMEAU: Immunité
Untergeordnet	7 Datensätze 1. *Angeborene Immunität* *Immunität <Medizin>* 2. *Autoimmunität* *Immunität <Medizin>* 3. ...
Thema in	6129 Publikationen 1. *Geben Sie Bakterien und Viren keine Chance* *Krüger, Karsten. - München : Scorpio, [2019]* 2. *Geben Sie Bakterien und Viren keine Chance!* *Krüger, Karsten. - München : Scorpio Verlag, 2019, 1. Auflage* 3. ...

Рис. 7. «Иммунитет» в электронном каталоге Национальной библиотеки Германии (с указанием иерархических отношений и связей с индексами УДК)

элементами других LOD-словарей, выражающими ту же тему (концепт).

На рис. 5 видны связи с вышестоящими и нижестоящими понятиями и работающие связи (ссылки) с теми же понятиями в шести других LOD-словарях. При этом указаны типы связей между темами: *Exact Matching (точное соответствие) и Clousely Matching (близкое по значению)*.

По связям, опубликованным на рассмотренном сайте Библиотеки Конгресса США (рис. 5), переходим к соответствующим страницам других словарей (рис. 6 и 7).

Пример 7. В РГБ был проведен эксперимент по установлению семантического соответствия между индексами УДК и полного варианта таблиц ББК. На рис. 8 показана страница представления индекса УДК *76* в LOD (вышестоящий и нижестоящие уровни, а также ссылки на ассоциативные разделы), на рис. 9 — страница с простым примером установления его соответствия с индексом ББК *Щ15*. Большинство соответствий между другими индексами этих классификаций выглядят далеко не так просто.

Таким образом, между URI индекса УДК *76* и URI индекса ББК *Щ15* устанавливается точное соответствие *skos:exactMatch*, что дает основание в будущем реализовать программный переход при поиске по индексу ББК в каталог, использующий УДК, или наоборот. Аналогичная возможность откроется, например, для пересылки запроса по индексу ББК *Щ15* в нормативную базу данных FAST Linked Data [43] с целью поиска в каталоге WordCat OCLC по теме «*Art. Graf-ic*» и далее в другие ресурсы, связанные с FAST (см. раздел *Sources and other links*, рис. 10).

В последнее время РГБ разрабатывает проект установления соответствий между концептами основных таблиц полного и среднего [38]

UDC Summary Linked Data Русский (Russian)

URI	http://udcdata.info/065294 RDF
Notation	76
Caption	Графические искусства. Графика. Гравюра
See also	655 Полиграфическая промышленность. Издательское дело. Книжная торговля 681.6 Множительные и печатные машины и оборудование
Broader class	7 ИСКУССТВО. РАЗВЛЕЧЕНИЯ. ЗРЕЛИЩА. СПОРТ
Narrower classes	761 Гравюра высокой печати (выпуклая) 762 Гравюра глубокой печати (углубленная) 763 Гравюра плоской печати. Литография 766 Прикладная графика

The UDC Summary (UDCS) provides a selection of around 2,400 classes from the whole scheme which comprises more than 69,000 entries. Please send questions and suggestions to udcs@udcc.org

UDC CONSORTIUM

The data provided in this Summary is released under the Creative Commons Attribution Share Alike 3.0 license [more]

Рис. 8. Раздел 76 УДК «Графические искусства. Графика. Гравюра» в LOD (версия на русском языке)

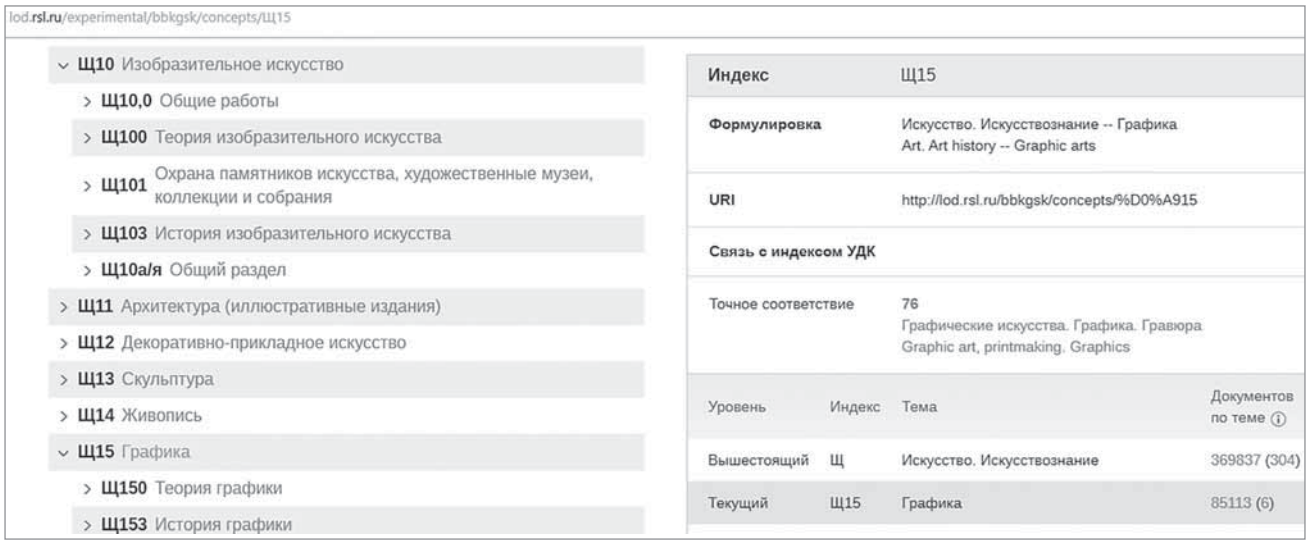


Рис. 9. Экспериментальное установление связи раздела 76 УДК «Графические искусства. Графика. Гравюра» и раздела ББК Щ15 «Искусство. Искусствознание -- Графика» в среде LOD

вариантов таблиц ББК. Это в определенной степени позволит соединить в едином семантическом пространстве поисковые средства библиотек различного уровня. Если соединить URI индекса среднего варианта таблиц *85.15 «Искусство. Искусствознание -- Изобразительное искусство и архитектура -- Графика»* с URI индекса из полных таблиц *Щ15 «Искусство. Искусствознание -- Графика»* связью *skos:exactMatch,* то в дальнейшем на базе языка запросов SPARQL можно строить маршруты поиска по данному индексу и в рассмотренных выше LOD-словарях.

Также следует отметить важное обстоятельство: для представления в Semantic Web нормативных/авторитетных данных опубликованы пространства имен для моделей семейства FR [44], FRBR [45], FRAD [46; 47] и FRSAD [48; 49]. Это открывает широкие возможности представления и связей в LOD такого рода библиотечных данных с элементами Классификационной системы.

Graphic arts Find in WorldCat
http://fast.oclc.org/searchfast/?&limit=keywords&facet=all&query=graphic&sort=usage+desc&start=0#&single=fst00946595&fullview=simple&sep=click
USED FOR:
- Art, Graphic
- Arts, Graphic
- Graphic design (Graphic arts)
- Graphics
- SEE ALSO:
- Art(OCoLC)fst00815177
- Visual communication(OCoLC)fst01167997
- USAGE:
- LC (2017) Subject Usage: 968 (1,282)
- WC (2017) Subject Usage: 23,871 (28,993)
- RECORD ID:
- fst00946595
- SOURCES AND OTHER LINKS:
- Graphic arts--(DLC)sh 85056474
- Graphic arts--http://en.wikipedia.org/wiki/Graphic_arts--https://www.wikidata.org/wiki/Q911059
- Graphics--http://en.wikipedia.org/wiki/Graphics--https://www.wikidata.org/wiki/Q1027879
- Arts graphiques--http://data.bnf.fr/ark:/12148/cb11931640m
- Grafik--http://d-nb.info/gnd/4021845-4

Рис. 10. Представление связей концепта «Искусство, графика» в FAST с такими же по смыслу концептами в других LOD-словарях

Перспективы

РГБ занимается проблемами представления знаний в библиотеках не одно десятилетие, регулярно освещая свою деятельность в публикациях, упомянутых выше, к которым полезно добавить ссылки на статьи по вопросам семантического поиска на основе классификации [50; 51], созданию систем организации знаний [52; 53] и представлению классификации в Семантической паутине [54].

Пришла пора международного сотрудничества российских библиотек в Semantic Web. В настоящее время РГБ предлагает нашему библиотечному сообществу заняться формированием в России единой модели пространства знаний, основанной на действующей Классификационной системе организации знаний и LOD-представлениях наиболее используемых в библиотеках словарей (Среднего варианта таблиц ББК, таблиц УДК, систем предметных рубрик, тезаурусов, нормативных данных на имена лиц, названия организаций и географические названия) [55]. Чашечка отличного кофе не заставит себя ждать во время совместной работы специалистов различных библиотек в гостеприимных стенах РГБ.

Список источников

1. *Лаврёнова О.А.* Технологии открытых связанных данных и «дорожные карты» как навигаторы для пользователей библиотек // Информационные ресурсы — футурологический аспект: планы, прогнозы, перспективы : материалы X Всерос. науч.-практ. конф. «Электронные ресурсы библиотек, музеев, архивов» (30—31 окт., 2014 г., Санкт-Петербург). Санкт-Петербург : Политехника-сервис, 2014. С. 146—154.
2. *Бернерс-Ли Т., Хендлер Дж., Лассила О.* Семантическая Сеть [Электронный ресурс]. URL: http://www.cs.man.ac.uk/~ezolin/logic/semantic_web_utf.html (дата обращения: 25.02.2020).
3. *Михайлов А.И., Черный А.И., Гиляревский Р.С.* Основы информатики. Москва : Наука, 1968. 757 с.
4. *Гиляревский Р.С.* Основы информатики : курс лекций [Электронный ресурс]. Москва, 2003. URL: http://www.hse.ru/data/2012/04/10/1251615157/БИ_курс%20лекций.pdf (дата обращения: 25.02.2020).
5. Networked Knowledge Organization. Systems/Services/Structures. NKOS [Электронный ресурс]. URL: http://nkos.slis.kent.edu (дата обращения: 25.02.2020).
6. NKOS — 15th European Networked Knowledge Organization Systems [Электронный ресурс] // 20th International Conference on Theory and Practice of Digital Libraries. TPDL 2016. URL: www.tpdl2016.org/nkos (дата обращения: 25.02.2020).
7. *Mayr P., Tudhope D., Clarke S.D. et al.* Recent Applications of Knowledge Organization Systems: Introduction to a Special Issue // International Journal on Digital Libraries. 2016. Vol. 17, № 1. P. 1—4.
8. *Лаврёнова О.А.* Тематический поиск в электронных каталогах и электронных библиотеках // Библиотековедение. 2004. № 5. С. 42—50.
9. *Лаврёнова О.А.* Методика разработки информационно-поискового тезауруса. Москва : Пашков дом, 2001. 54 с. URL: http://dlib.rsl.ru/01000713291 (дата обращения: 25.02.2020).
10. *Гендина Н.И.* Информационно-поисковые тезаурусы: основные виды и области применения [Электронный ресурс]. URL: http://masters.donntu.org/2013/fknt/bilykn/library/8.htm (дата обращения: 25.02.2020).
11. Библиотечно-библиографическая классификация : таблицы для научных библиотек / Гос. б-ка СССР им. В.И. Ленина и др. ; редкол.: О.П. Тесленко (гл. ред.) [и др.]. Москва : [б. и.], 1960.
12. *Лаврёнова О.А.* Возможности пользователя при поиске информации в электронных библиотеках, или «Витязь на распутье» // Библиотековедение. 2013. № 3. С. 43—52. DOI: 10.25281/0869-608X-2013-0-3-43-52.
13. *Лаврёнова О.А.* Семантические средства библиографического поиска в Российской государственной библиотеке // Общетеоретические и футурологические проблемы библиографии. Библиографическая запись как основа формирования библиографических ресурсов : материалы II Междунар. библиогр. конгресса «Библиография: взгляд в будущее» (Москва, 6—8 окт. 2015 г.) / Рос. гос. б-ка. Москва : Пашков дом, 2016. С. 309—323.
14. Классификационная система [Электронный ресурс] // Linked Open Data. URL: https://lod.rsl.ru/bbkgsk/concepts/ (дата обращения: 25.02.2020).
15. О лицензиях [Электронный ресурс] // Creative Commons. URL: https://creativecommons.org/licenses/?lang=ru (дата обращения: 25.02.2020).
16. Семантическая паутина [Электронный ресурс] // Википедия. URL: https://ru.wikipedia.org/wiki/Семантическая_паутина (дата обращения: 25.02.2020).
17. *Zeng M.L., Mayr P.* Knowledge Organization Systems (KOS) in the Semantic Web : Amulti-Dimensional Review // International Journal on Digital Libraries. 2019. Vol. 20, № 3. P. 209—230.
18. *Binding C., Tudhope D.* Improving Interoperability Using Vocabulary Linked Data // International Journal on Digital Libraries. 2016. Vol. 17, № 1. P. 5—21.

19. Namespaces in XML 1.0. 3rd Edition [Электронный ресурс] // W3C Recommendation. URL: https://www.w3.org/TR/xml-names/ (дата обращения: 25.02.2020).
20. Namespace [Электронный ресурс] // Wikipedia. URL: https://en.wikipedia.org/wiki/Namespace (дата обращения: 25.02.2020).
21. *Серебряков В.А., Шорин О.Н.* Проблемы семантической интеграции библиотечных данных // Библиотековедение. 2014. № 5. С. 41—47. DOI: 10.25281/0869-608X-2014-0-5-41-47.
22. BIBFRAME. Bibliographic Framework Initiative [Электронный ресурс] // The Library of Congress. URL: https://www.loc.gov/bibframe/ (дата обращения: 01.04.2020).
23. Linked Open Vocabularies (LOV) [Электронный ресурс]. URL: http://lov.okfn.org/dataset/lov (дата обращения: 25.02.2020).
24. *Lavrenova O., Pavlov V.* The Publication of a Knowledge Organization Classification System as a Linked Data Vocabulary [Электронный ресурс] // Proceedings of the 17th European Networked Knowledge Organization Systems Workshop Co-Located with the 21st International Conference on Theory and Practice of Digital Libraries 2017 (TPDL 2017), (21st September, 2017, Thessaloniki, Greece). CEUR Workshop Proceedings. Vol. 1937. P. 62—73. URL: http://ceur-ws.org/Vol-1937/paper6.pdf (дата обращения: 25.02.2020).
25. *Christoph P.* Building a High Performance Environment for RDF Publishing [Электронный ресурс] // Semantic Web in Libraries. SWIB12 (26—28 November 2012, Cologne, Germany). URL: http://swib.org/swib12/slides/Christoph_SWIB12_117.pdf (дата обращения: 25.02.2020).
26. *Лаврёнова О.А., Павлов В.В.* Библиотечно-библиографическая классификация как традиционная система организации знаний в среде открытых связанных данных // Научные и технические библиотеки. 2017. № 4. С. 44—60. DOI: 10.33186/1027-3689-2017-4-44-60.
27. URI [Электронный ресурс] // Википедия. URL: https://ru.wikipedia.org/wiki/URI (дата обращения: 25.02.2020).
28. RDF 1.1 Concepts and Abstract Syntax [Электронный ресурс] // W3C Recommendation. URL: https://www.w3.org/TR/rdf11-concepts/ (дата обращения: 25.02.2020).
29. Resource Description Framework [Электронный ресурс] // Википедия. URL: https://ru.wikipedia.org/wiki/Resource_Description_Framework (дата обращения: 25.02.2020).
30. *Жлобинская О.Н.* Как представлять библиографические данные в RDF [Электронный ресурс] // RUSMARC. URL: http://www.rusmarc.ru/publish/bibdateRDF.pdf (дата обращения: 25.02.2020).
31. *Лаврёнова О.А.* Опыт представления классификационных метаданных в среде Semantic Web для поиска электронных ресурсов библиотек // Сборники Президентской библиотеки. Вып. 9 : Гуманитарные исследования и цифровая среда: наука и практика = Humanitarian Research and the Digital Environment: Theory and Practice : сб. науч. тр. Санкт-Петербург, 2019. С. 54—69. (Серия «Электронная библиотека»). URL: https://www.prlib.ru/item/1279738 (дата обращения: 25.02.2020).
32. SKOS Simple Knowledge Organization System Primer [Электронный ресурс] // W3C Working Group Note. URL: http://www.w3.org/TR/skos-primer/ (дата обращения: 25.02.2020).
33. ARC SPARQL+ Endpoint (v2011-12-01) [Электронный ресурс]. URL: http://metadataregistry.org/endpoint.php (дата обращения: 25.02.2020).
34. SPARQL Protocol for RDF/ [Электронный ресурс] // W3C Recommendation. URL: https://www.w3.org/TR/rdf-sparql-protocol/ (дата обращения: 25.02.2020).
35. Классификационная система [Электронный ресурс] // Linked Open Data. URL: https://lod.rsl.ru/#dataAccess (дата обращения: 25.02.2020).
36. *Лаврёнова О.А.* Тематический поиск электронных ресурсов на основе классификационной модели организации знаний // Университетская книга. 2017. № 9. С. 58—62.
37. *Шварцман М.Е., Найдин О.П.* Linked Open Data как средство обогащения поисковых запросов // Университетская книга. 2015. № 12. С. 66—71.
38. Библиотечно-библиографическая классификация [Электронный ресурс] // Российская государственная библиотека : офиц. сайт. URL: http://bbk.rsl.ru/external/bbk?block=ETALON (дата обращения: 25.02.2020).
39. UDC Summary Linked Data [Электронный ресурс]. URL: http://udcdata.info (дата обращения: 25.02.2020).
40. Справочник географических названий [Электронный ресурс] // Российская государственная библиотека : офиц. сайт. URL: http://aleph.rsl.ru/F/?func=file&file_name=find-b&local_base=tst11 (дата обращения: 25.02.2020).
41. *Лаврёнова О.А.* Национальный файл географических названий — новый проект Российской государственной библиотеки // Библиотековедение. 2006. № 2. С. 46—53.
42. *Шамардина Л.А.* Использование тезауруса MeSH в лингвистическом обеспечении базы данных «MedArt» [Электронный ресурс] // ГПНТБ Рос-

сии. URL: http://www.gpntb.ru/libcom7/disk/21.pdf (дата обращения: 25.02.2020).
43. FAST (Faceted Application of Subject Terminology) [Электронный ресурс] // OCLC FAST. URL: https://fast.oclc.org (дата обращения: 25.02.2020).
44. Representing the FR Family in the Semantic Web // Cataloging & Classification Quarterly. 2012. Vol. 50, issue 5—7. P. 724—741. DOI: 10.1080/01639374.2012.679881.
45. Element Sets: Show Detail for FRBRer model [Электронный ресурс] // Open Metadata Registry. URL: http://metadataregistry.org/schema/show/id/5.html (дата обращения: 25.02.2020).
46. Element Sets: Show Detail for FRAD model [Электронный ресурс] // Open Metadata Registry. URL: http://metadataregistry.org/schema/show/id/24.html (дата обращения: 25.02.2020).
47. Функциональные требования к авторитетным данным : концептуальная модель : заключительный отчет, декабрь 2008 / под ред. Г.Е. Патона ; Рабочая группа ИФЛА по разработке функциональных требований к авторитетным записям и их нумерации (FRANAR) ; [пер. с англ. О.А. Лаврёнова]. Санкт-Петербург, 2011. 115 с. [Электронный ресурс] // The International Federation of Library Associations and Institutions (IFLA). URL: https://www.ifla.org/files/assets/cataloguing/frad/frad_2009-ru.pdf (дата обращения: 25.02.2020).
48. Element Sets: Show Detail for FRSAD model [Электронный ресурс] // Open Metadata Registry. URL: http://metadataregistry.org/schema/show/id/26.html (дата обращения: 25.02.2020).
49. Функциональные требования к предметным авторитетным данным (FRSAD) : концептуальная модель / под ред. М.Л. Зенг, М. Жумер, А. Салаба ; [пер. с англ. А.В. Хлебаевой под ред. Е.И. Загорской, Н.П. Никольцевой] ; Междунар. федерация библ. ассоциаций и учреждений, Рос. нац. б-ка. Санкт-Петербург : РНБ, 2016. 96 с.
50. *Lavrenova O.* Thematical Search in the Automated Library System on the Basis of the Library Bibliographical Classification (LBC) with Subject Entry // 57th IFLA General Conference (18—24 August, 1991, Moscow). Paper 4-31-36.
51. *Lavrenova O.A.* The National Library Bibliographic Classification (BBK) as a Basis for Subject Search in the Integrated RSL Digital Library : The Project Presentation [Электронный ресурс] // Networked Knowledge Organization Systems and Services : The 11th European Networked Knowledge Organization Systems (NKOS) Workshop. Workshop at the TPDL Conference, Paphos, Cyprus, 2012. URL: https://nkos-eu.github.io/2012/presentations/TPDL-2012-Lavrenova-w.pdf (дата обращения: 25.02.2020).
52. *Lavrenova O.* Knowledge Organization Levels for Computerized Information Systems // Knowledge Organization. 1993. Vol. 20, № 4. P. 207.
53. *Винберг А.А., Лаврёнова О.А.* Эффективное использование моделей организации знаний для тематического поиска в электронных ресурсах [Электронный ресурс]. Science Online XXII. Электронные информационные ресурсы для науки и образования (26 мая — 2 июня 2019 г., Испания, остров Майорка) // Научная электронная библиотека eLIBRARY.RU. URL: https://www.elibrary.ru/projects/conference/majorca2019/presentations/Vinberg-Lavrenova.pptx (дата обращения: 25.02.2020).
54. *Лаврёнова О.А.* Представление классификации в Семантической паутине // Информационное обслуживание в век электронных коммуникаций : XI Всерос. науч.-практ. конф. «Электронные ресурсы библиотек, музеев, архивов» (2—3 ноября 2016 г., Санкт-Петербург) : сб. материалов. Санкт-Петербург, 2016. С. 73—92.
55. *Лаврёнова О.А.* Задача интеграции технологий тематического поиска библиотечных ресурсов на основе словарей связанных данных (LOD vocabularies) // Румянцевские чтения — 2019 : материалы Междунар. науч.-практ. конф. (23—24 апреля 2019 г.). Москва : Пашков дом, 2019. Ч. 2. С. 104—109.

Иллюстративный материал предоставлен авторами статьи

Modern Users of Libraries and the Linked Open Data Environment

Olga A. Lavrenova [a] *, **Andrey A. Vinberg** **,
Russian State Library, 3/5 Vozdvizhenka Str., Moscow, 119019, Russia
[a] ORCID 0000-0002-4779-9222
E-mail: * LavrenovaOA@rsl.ru, ** VinbergAA@rsl.ru

Abstract. The goal of any library is to ensure high quality and general availability of information retrieval tools. The paper describes the project implemented by the Russian State Library (RSL) to present Library Bibliographic Classification as a Networked Knowledge Organization System. The project goal is to support content and provide tools for ensuring system's interoperability with other resources of the same nature (i.e. with Linked Data Vocabularies) in the global network environment. The project was partially supported by the Russian Foundation for Basic Research (RFBR).
The RSL General Classified Catalogue (GCC) was selected as the main data source for the Classification system of knowledge organization. The meaning of each classification number is expressed by complete string of wordings (captions), rather than the last level caption alone. Data converted to the Resource Description Framework (RDF) files based on the standard set of properties defined in the Simple Knowledge Organization System (SKOS) model was loaded into the semantic storage for subsequent data processing using the SPARQL query language. In order to enrich user queries for search of resources, the RSL has published its Classification System in the form of Linked Open Data (https://lod.rsl.ru) for searching in the RSL electronic catalogue. Currently, the work is underway to enable its smooth integration with other LOD vocabularies. The SKOS mapping tags are used to differentiate the types of connections between SKOS elements (concepts) existing in different concept schemes, for example, UDC, MeSH, authority data.
The conceptual schemes of the leading classifications are fundamentally different from each other. Establishing correspondence between concepts is possible only on the basis of lexical and structural analysis to compute the concept similarity as a combination of attributes.
The authors are looking forward to working with libraries in Russia and other countries to create a common space of Linked Open Data vocabularies.

Key words: Linked Open Data, Linked Data vocabularies, Network Knowledge Organization Systems, NKOS, Semantic Web, Library Bibliographic Classification, information retrieval, semantic road maps, user's search intentions, SKOS, RDF, virtual systematic catalogue, query enrichment, electronic libraries, library space organization.

Citation: Lavrenova O.A., Vinberg A.A. Modern Users of Libraries and the Linked Open Data Environment, *Bibliotekovedenie* [Russian Journal of Library Science], 2020, vol. 69, no. 3, pp. 243—260. DOI: 10.25281/0869-608X2020-69-3-243-260.

References

1. Lavrenova O.A. The Open Linked Data Technologies and "Road Maps" as Navigators for Library Users, *Informatsionnye resursy — futurologicheskii aspekt: plany, prognozy, perspektivy: materialy X Vseros. nauch.-prakt. konf. "Elektronnye resursy bibliotek, muzeev, arkhivov" (30—31 oktyabrya, 2014 g., Sankt-Peterburg)* [Information Resources — Futurological Aspect: Plans, Forecasts, Prospects: Proceedings of the 10th All-Russian Sci.-Pract. Conf. "Electronic Resources of Libraries, Museums, Archives" (October 30—31, 2014, Saint Petersburg)]. St. Petersburg, Politekhnika-servis Publ., 2014, pp. 146—154 (in Russ.).
2. Berners-Lee T., Hendler J., Lassila O. *The Semantic Web*. Available at: http://www.cs.man.

ac.uk/~ezolin/logic/semantic_web_utf.html (accessed 25.02.2020) (in Russ.).
3. Mikhailov A.I., Chernyi A.I., Gilyarevsky R.S. *Osnovy informatiki* [The Basics of Informatics]. Moscow, Nauka Publ., 1968, 757 p.
4. Gilyarevsky R.S. *Osnovy informatiki: kurs lektsii* [The Basics of Informatics: lecture course]. Moscow, 2003. Available at: https://www.hse.ru/data/2012/04/10/1251615157/%D0%91%D0%98_%D0%BA%D1%83%D1%80%D1%81%20%D0%BB%D0%B5%D0%BA%D1%86%D0%B8%D0%B9.pdf (accessed 25.02.2020).
5. *Networked Knowledge Organization. Systems/Services/Structures. NKOS*. Available at: http://nkos.slis.kent.edu (accessed 25.02.2020).
6. NKOS — 15th European Networked Knowledge Organization Systems, *20th International Conference on Theory and Practice of Digital Libraries. TPDL 2016*. Available at: www.tpdl2016.org/nkos (accessed 25.02.2020).
7. Mayr P., Tudhope D., Clarke S.D. et al. Recent Applications of Knowledge Organization Systems: Introduction to a Special Issue, *International Journal on Digital Libraries*, 2016, vol. 17, no. 1, pp. 1—4.
8. Lavrenova O.A. Subject Search in Digital Catalogues and Libraries, *Bibliotekovedenie* [Russian Journal of Library Science], 2004, no. 5, pp. 42—50 (in Russ.).
9. Lavrenova O.A. *Metodika razrabotki informatsionno-poiskovogo tezaurusa* [Methodology for Developing an Information and Search Thesaurus]. Moscow, Pashkov Dom Publ., 2001, 54 p. Available at: http://dlib.rsl.ru/01000713291 (accessed 25.02.2020).
10. Gendina N.I. *Informatsionno-poiskovye tezaurusy: osnovnye vidy i oblasti primeneniya* [Information and Search Thesauri: Main Types and Applications]. Available at: http://masters.donntu.org/2013/fknt/bilykn/library/8.htm (accessed 25.02.2020).
11. Teslenko O.P. (ed.) *Bibliotechno-bibliograficheskaya klassifikatsiya: tablitsy dlya nauchnykh bibliotek* [Library-Bibliographical Classification: Schedules for Scientific Libraries]. Moscow, 1960.
12. Lavrenova O.A. User's Options for Information Search in the Electronic Libraries, or "Knight at the Crossroads", *Bibliotekovedenie* [Russian Journal of Library Science], 2013, no. 3, pp. 43—52. DOI: 10.25281/0869-608X-2013-0-3-43-52 (in Russ.).
13. Lavrenova O.A. Semantic Means of Bibliographic Search in the Russian State Library, *Obshcheteoreticheskie i futurologicheskie problemy bibliografii. Bibliograficheskaya zapis' kak osnova formirovaniya bibliograficheskikh resursov: materialy II Mezhdunar. bibliogr. kongressa "Bibliografiya: vzglyad v budushchee" (Moskva, 6—8 oktyabrya 2015 g.)* [General Theoretical and Futurological Issues of Bibliography. Bibliographic Record as a Basis for Bibliographic Resources Formation: Proceedings of the 2nd Int. Bibliog. Congress "Bibliography: Looking into the Future" (Moscow, October 6—8, 2015)]. Moscow, Pashkov Dom Publ., 2016, pp. 309—323 (in Russ.).
14. Classification System, *Linked Open Data*. Available at: https://lod.rsl.ru/bbkgsk/concepts/ (accessed 25.02.2020) (in Russ.).
15. About the Licenses, *Creative Commons*. Available at: https://creativecommons.org/licenses/?lang=ru (accessed 25.02.2020).
16. Semantic Web, *Wikipedia*. Available at: https://en.wikipedia.org/wiki/Semantic_Web (accessed 25.02.2020) (in Russ.).
17. Zeng M.L., Mayr P. Knowledge Organization Systems (KOS) in the Semantic Web: Amulti-Dimensional Review, *International Journal on Digital Libraries*, 2019, vol. 20, no. 3, pp. 209—230.
18. Binding C., Tudhope D. Improving Interoperability Using Vocabulary Linked Data, *International Journal on Digital Libraries*, 2016, vol. 17, no. 1, pp. 5—21.
19. Namespaces in XML 1.0. 3rd Edition, *W3C Recommendation*. Available at: https://www.w3.org/TR/xml-names/ (accessed 25.02.2020).
20. Namespace, *Wikipedia*. Available at: https://en.wikipedia.org/wiki/Namespace (accessed 25.02.2020).
21. Serebryakov V.A., Shorin O.N. Problems of Semantic Integration of Library Data, *Bibliotekovedenie* [Russian Journal of Library Science], 2014, no. 5, pp. 41—47. DOI: 10.25281/0869-608X-2014-0-5-41-47 (in Russ.).
22. BIBFRAME. Bibliographic Framework Initiative, *The Library of Congress*. Available at: https://www.loc.gov/bibframe/ (accessed 01.04.2020).
23. *Linked Open Vocabularies (LOV)*. Available at: http://lov.okfn.org/dataset/lov (accessed 25.02.2020).
24. Lavrenova O., Pavlov V. The Publication of a Knowledge Organization Classification System as a Linked Data Vocabulary, *Proceedings of the 17th European Networked Knowledge Organization Systems Workshop Co-Located with the 21st International Conference on Theory and Practice of Digital Libraries 2017 (TPDL 2017), (21st September, 2017, Thessaloniki, Greece). CEUR Workshop Proceedings*, vol. 1937, pp. 62—73. Available at: http://ceur-ws.org/Vol-1937/paper6.pdf (accessed 25.02.2020).
25. Christoph P. Building a High Performance Environment for RDF Publishing, *Semantic Web in Libraries. SWIB12 (26—28 November 2012, Cologne, Germany)*. Available at: http://swib.org/swib12/slides/Christoph_SWIB12_117.pdf (accessed 25.02.2020).

26. Lavrenova O.A., Pavlov V.V. Library Bibliographic Classification as a Traditional Knowledge Organization System in the Linked Open Data Environment, *Nauchnye i tekhnicheskie biblioteki* [Scientific and Technical Libraries], 2017, no. 4, pp. 44—60 (in Russ.). DOI: 10.33186/1027-3689-2017-4-44-60.
27. URI, *Wikipedia*. Available at: https://en.wikipedia.org/wiki/Uniform_Resource_Identifier (accessed 25.02.2020) (in Russ.).
28. RDF 1.1 Concepts and Abstract Syntax, *W3C Recommendation*. Available at: https://www.w3.org/TR/rdf11-concepts/ (accessed 25.02.2020).
29. Resource Description Framework, *Wikipedia*. Available at: https://en.wikipedia.org/wiki/Resource_Description_Framework (accessed 25.02.2020) (in Russ.).
30. Zhlobinskaya O.N. How to Present Bibliographic Data in the RDF, *RUSMARC*. Available at: http://www.rusmarc.ru/publish/bibdateRDF.pdf (accessed 25.02.2020) (in Russ.).
31. Lavrenova O.A. An Experience in Presenting Classification Metadata in the Semantic Web Environment for Searching Electronic Resources of Libraries, *Sborniki Prezidentskoi biblioteki. Vyp. 9: Gumanitarnye issledovaniya i tsifrovaya sreda: nauka i praktika: sb. nauch. tr.* [Presidential Library's Collections. Issue 9. Humanitarian Research and the Digital Environment: Theory and Practice: collected scientific papers]. St. Petersburg, 2019, pp. 54—69. Available at: https://www.prlib.ru/item/1279738 (accessed 25.02.2020) (in Russ.).
32. SKOS Simple Knowledge Organization System Primer, *W3C Working Group Note*. Available at: http://www.w3.org/TR/skos-primer/ (accessed 25.02.2020).
33. *ARC SPARQL+ Endpoint (v2011-12-01)*. Available at: http://metadataregistry.org/endpoint.php (accessed 25.02.2020).
34. SPARQL Protocol for RDF, *W3C Recommendation*. Available at: https://www.w3.org/TR/rdf-sparql-protocol/ (accessed 25.02.2020).
35. Classification System, *Linked Open Data of the Russian State Library*. Available at: https://lod.rsl.ru/#dataAccess (accessed 25.02.2020) (in Russ.).
36. Lavrenova O.A. Subject Search for Electronic Resources Based on the Classification Model of Knowledge Organization, *Universitetskaya kniga* [University Book], 2017, no. 9, pp. 58—62 (in Russ.).
37. Shvartsman M.E., Naidin O.P. Linked Open Data as a Means of Enriching Search Queries, *Universitetskaya kniga* [University Book], 2015, no. 12, pp. 66—71 (in Russ.).
38. Library-Bibliographical Classification, *Rossiiskaya gosudarstvennaya biblioteka: Ofits. sait* [Russian State Library: official website]. Available at: http://bbk.rsl.ru/external/bbk?block=ETALON (accessed 25.02.2020) (in Russ.).
39. *UDC Summary Linked Data*. Available at: http://udcdata.info (accessed 25.02.2020).
40. Reference Book of Geographical Names, *Rossiiskaya gosudarstvennaya biblioteka: Ofits. sait* [Russian State Library: official website]. Available at: http://aleph.rsl.ru/F/?func=file&file_name=find-b&local_base=tst11 (accessed 25.02.2020) (in Russ.).
41. Lavrenova O.A. National File of Geographic Names — a New Project of Russian State Library, *Bibliotekovedenie* [Russian Journal of Library Science], 2006, no. 2, pp. 46—53 (in Russ.).
42. Shamardina L.A. Using MeSH Thesaurus in the Linguistic Support of the "MedArt" Database, *GPNTB Rossii* [Russian National Public Library for Science and Technology]. Available at: http://www.gpntb.ru/libcom7/disk/21.pdf (accessed 25.02.2020) (in Russ.).
43. FAST (Faceted Application of Subject Terminology), *OCLC FAST*. Available at: https://fast.oclc.org (accessed 25.02.2020).
44. Representing the FR Family in the Semantic Web, *Cataloging & Classification Quarterly*, 2012, vol. 50, issue 5—7, p. 724—741. DOI: 10.1080/01639374.2012.679881.
45. Element Sets: Show Detail for FRBRer model, *Open Metadata Registry*. Available at: http://metadataregistry.org/schema/show/id/5.html (accessed 25.02.2020).
46. Element Sets: Show Detail for FRAD model, *Open Metadata Registry*. Available at: http://metadataregistry.org/schema/show/id/24.html (accessed 25.02.2020).
47. Patton G.E. (ed.) Functional Requirements of Authority Data: A Conceptual Model: Final Report, December 2008, *The International Federation of Library Associations and Institutions (IFLA)*. Available at: https://www.ifla.org/files/assets/cataloguing/frad/frad_2009-ru.pdf (accessed 25.02.2020).
48. Element Sets: Show Detail for FRSAD model, *Open Metadata Registry*. Available at: http://metadataregistry.org/schema/show/id/26.html (accessed 25.02.2020).
49. Zeng M.L., Žumer M., Salaba A. (eds). *Functional Requirements for Subject Authority Data (FRSAD): A Conceptual Model*. St. Petersburg, RNB Publ., 2016, 96 p. (in Russ.).
50. Lavrenova O. Thematical Search in the Automated Library System on the Basis of the Library Bibliographical Classification (LBC) with Subject Entry, *57th IFLA General Conference (18—24 August, 1991, Moscow)*, paper 4-31-36.
51. Lavrenova O.A. The National Library Bibliographic Classification (BBK) as a Basis for Subject Search in the Integrated RSL Digital Library: The Project Pre-

sentation, *Networked Knowledge Organization Systems and Services: The 11th European Networked Knowledge Organization Systems (NKOS) Workshop. Workshop at the TPDL Conference, Paphos, Cyprus, 2012*. Available at: https://nkos-eu.github.io/2012/presentations/TPDL-2012-Lavrenova-w.pdf (accessed 25.02.2020).

52. Lavrenova O. Knowledge Organization Levels for Computerized Information Systems, *Knowledge Organization*, 1993, vol. 20, no. 4, p. 207.
53. Vinberg A.A., Lavrenova O.A. Effective Use of Knowledge Organization Models for Subject Search in Electronic Resources. Science Online XXII. Electronic Information Resources for Science and Education (May 26 — June 2, 2019, Majorca, Spain), *Nauchnaya elektronnaya biblioteka eLIBRARY.RU* [Scientific Electronic Library eLIBRARY.RU]. Available at: https://www.elibrary.ru/projects/conference/majorca2019/presentations/Vinberg-Lavrenova.pptx (accessed 25.02.2020) (in Russ.).
54. Lavrenova O.A. Representation of Classification in the Semantic Web, *Informatsionnoe obsluzhivanie v vek elektronnykh kommunikatsii: XI Vseros. nauch.-prakt. konf. "Elektronnye resursy bibliotek, muzeev, arkhivov" (2—3 noyabrya 2016 g., Sankt-Peterburg): sb. materialov* [Information Service in the Age of Electronic Communications: Proceedings of the 11th All-Russian Sci.-Pract. Conf. "Electronic Resources of Libraries, Museums, Archives" (November 2—3, 2016, Saint Petersburg)]. St. Petersburg, 2016, pp. 73—92 (in Russ.).
55. Lavrenova O.A. The Technology Integration for Subject Search of Library Resources Based on LOD Vocabularies, *Rumyantsevskie chteniya — 2019: materialy Mezhdunar. nauch.-prakt. konf. (23—24 aprelya 2019 g.)* [Proceedings of the Int. Sci.-Pract. Conf. "Rumyantsev Readings — 2019" (April 23—24, 2019)]. Moscow, Pashkov Dom Publ., 2019, part 2, pp. 104—109 (in Russ.).

Анонс

Книги на карантине. Правила обращения с документами

На сайте Российской государственной библиотеки (РГБ) собраны рекомендации для библиотек по работе с книгами в условиях борьбы с коронавирусом. В текущей ситуации важно не только соблюдать правила дезинфекции, но и сохранить целостность документов.

Регулярно обновляющийся перечень рекомендаций доступен по адресу:
https://www.rsl.ru/ru/all-news/knigi-na-karantine-pravila

А.А. Кащеев, заведующий сектором превентивной консервации отдела реставрации библиотечных фондов РГБ, провел вебинар «Как изменится жизнь книг после карантина: что важно, а что делать нельзя», на котором были подняты следующие вопросы:

- проблема дезинфекции в консервации документов;
- почему кварцевание и дезинфекция может быть опасной для книг;
- советы профессионалам по обращению с книгами во время выхода из карантина. Вместе с экспертом мы разберемся как соблюсти грань между безопасностью читателей и сохранностью книг.

Видеозапись вебинара доступна на канале РГБ в YouTube:
http://www.youtube.com/watch?v=QLU-IA56gl4