

Место и роль лексикографического обеспечения в информационном поиске с применением иерархических классификаций

Рассматриваются роль и место лексикографического обеспечения в информационном поиске с применением иерархических классификаций на основе системного подхода и современных модельных представлений. Затронуты вопросы практического применения словарного лексикографического арсенала в процессе информационного поиска в Едином цифровом пространстве научного знания. Обсуждаются источники расширения лексикографического многообразия и особенности их интеграции применительно к проблеме информационного поиска. Показано, что расширение лексикографического потенциала полей описания понятий (классов), в основном, ориентировано на тематический поиск информации. Исключение составляет применение в информационном поиске таблиц сопоставления, которые могут быть направлены на путь информационного поиска по классификационным индексам.

Ключевые слова: информационный поиск, таблицы сопоставления, единое цифровое пространство

DOI: 10.36535/0548-0027-2022-03-3

ВВЕДЕНИЕ

Современное обеспечение эффективного информационного поиска направлено на новый уровень использования технологий обработки информации, в первую очередь – на переход к семантически значимому представлению научных знаний, извлекаемых из цифровой среды [1]. Исследователи всё больше склоняются к тому, что этим представлением являются иерархические структуры, способные контролировать данные не только в пределах дерева «сверху вниз», но и по уровням информационной значимости – по горизонтали. Последнее свойство подчеркивает важность самих иерархических структур как среды информационного поиска в Едином цифровом пространстве научных знаний [2].

Лексикографическая обеспеченность информационного поиска сегодня приобретает особое значение в связи с выделением Единого информационного пространства знаний и перемещением поиска в среду Интернет. Лексикографические методы основаны на «планомерной инвентаризации единиц языка посредством их лексикографирования». Это обеспечивается универсальностью методов и максимально полным охватом языковых средств, а также многоаспектностью анализа. Лексикографическая направленность релевантного информационного поиска ставит вопросы о роли, условиях и возможных ограничениях в применении этих методов. Для ответа на эти вопросы были рассмотрены модельные представления, отражающие сущность современного информационного поиска.

МОДЕЛЬ ИНФОРМАЦИОННОГО ПОИСКА В ЭЛЕКТРОННОМ КАТАЛОГЕ

На сегодняшний день библиотечным сообществом достигнута определённая ясность по исследованиям условий информационного поиска в электронных каталогах библиотек. Мнения специалистов уже сложились в круг представлений, которые расходятся, главным образом, в деталях, касающихся отдельных методических предпочтений [3, 4]. Эти представления об информационном поиске с участием классификационных иерархических информационно-библиографических схем для электронных каталогов библиотек в контексте системного анализа представляют собой трёхкомпонентную схему (рис. 1).

В основании схемы (см. рис. 1) располагаются элементы иерархических классификаций: классификационный индекс и описание класса, являющиеся необходимыми параметрами поиска, а вершину занимает компонент, на который обращён поиск, – библиографическая запись. Устойчивые связи между этими компонентами делают возможной организацию поиска информации. Модель информационного поиска выделяет два типа связей и, соответственно, два пути: по описанию классов (тематический) и по классификационным индексам (семантический).

Рассмотрим составляющие предлагаемой нами модели информационного поиска и их взаимосвязи более детально.

Описание класса – содержит понятия, отражённые в классификации. Этот компонент – многословное, нередко многострочное изложение сущности

понятий, заключённых в применяемой классификации. Главным требованием к конкретному описанию является строгое соответствие обсуждаемой теме и отсутствие семантических разрывов (целостность). Под целостностью мы понимаем возможность семантического центрирования конкретного описания в рамках темы, которой посвящён класс.

Семантическое центрирование – это антоним ризома (*rhizome*) Делёза [5], когда множество понятий образует цепочку по правилам конъюнкции (.../ и .../ и .../) и не характеризуется смысловым единством. Семантическое центрирование тем более актуально, что неполнота изложения понятий в описаниях классов часто вынуждает расширять эти описания, чтобы ликвидировать отставание представлений (понятий), размещённых в классификациях, от представлений современной науки.

При установлении целостности отдельные морфемы, лексемы и термины описания класса становятся объединённым общим смыслом лексикографическим арсеналом для релевантного информационного поиска. И чем более полно охарактеризовано понятие, тем более многообразен этот арсенал.

Индекс класса – это символичный указатель, создаваемый в соответствии с правилами, действующими в конкретной классификации. В иерархической структуре он служит для закрепления соподчинённости используемых в описаниях понятий на определенном иерархическом уровне и позволяет оценить их информационную значимость. Понятие (или тема) описания, сколь бы многословным оно ни было, семантически центрируется в классификационном индексе. В информационном поиске предназначение классификационного индекса – служить указателем для передачи единого смысла темы (понятия), к которой он приписан.

Классификационная диада: *описание класса – индекс класса* располагается в основании модели информационного поиска. Элементы этой диады имеют

функциональную семантическую связь, которая указывает на тождество индекса класса и понятия, отражённого в описании класса.

Библиографическая запись представляет собой подробное описание документа, включающее как его лексикографическое представление в виде библиографического описания, так и классификационный индекс, соотношённый с объектом, на который она обращена. Требование обязательности и правила индексирования нормированы и отражены в соответствующих стандартах [6-10]. Связь между индексированной библиографической записью и классификационным индексом также носит семантический характер и также образует диаду. Такая картина взаимоотношения библиографической записи с классификационным индексом свидетельствует о возможности выполнения семантического информационного поиска с помощью дорожной карты вида: *описание класса – индекс класса – библиографическая запись*. Устанавливая тождественность индекса классификации и индекса, присвоенного классифицируемому объекту, мы гарантируем, если не эквивалентность, то, как минимум, родство центрированных в них понятий. Дорожная карта вида: *описание класса – библиографическая запись* представляет собой путь *тематического информационного поиска*. На этом пути предоставляется возможность прямого диалога описания класса с библиографической записью. В таком диалоге выраженное вербальным языком описание класса имеет шанс выявить сходство своего лексикографического арсенала с арсеналом библиографического описания на уровне морфем, синонимов, терминов, фраз или ключевых слов, насколько ему это позволяет его лексикографический потенциал. При этом используются лингвистические связи между двумя фрагментами текста, представленными в нашем случае библиографическим описанием и описанием класса.



Рис. 1. Модель информационного поиска с использованием классификационных схем в информационно-библиографическом пространстве библиотечных электронных каталогов.



Рис. 2. Расширенная модель информационного поиска с использованием классификационных схем в условиях Единого цифрового пространства научных знаний.

Большинство библиотечных поисковых систем сегодня опираются, главным образом, на путь *тематического поиска информации* и не используют в полной мере возможности *семантического информационного поиска* по классификационным индексам.

С появлением в среде Интернета Единого цифрового информационно-библиографического пространства знаний время простых решений закончилось. В этом пространстве, составленном сотнями миллионов документов на всевозможных языках и индексируемых с применением различных классификаций, и сегодня царит библейское смешение языков. К сожалению, сегмент «индексируемого Интернета» сегодня представляет собой лишь лежащие на поверхности 5–10% общего объема информации, а сегмент «глубокого Интернета» уходит от вопросов унификации, становясь всё более недоступным. Однако при рассмотрении становится очевидным, что около 80% документов в доступном сегменте Интернета представляют собой систематизированный библиографический репертуар, имеющий классификационные индексы, а только около 20% документов лишены классификационных индексов. Сегмент «индексируемого Интернета» сегодня демонстрирует рост и это означает, что прежние модели, описывающие информационный поиск в электронных каталогах, не устарели, проблемы остались и настоятельно требуют решений с учётом новой реальности. Эта реальность сегодня обсуждается в проектах «семантического Web», связанных данных и прикладных задач, которые не могут быть решены без привлечения иерархических классификаций. Всё это требует нового взгляда на неполноту описания классов действующих классификаций и их соответствующего расширения. На рис. 2 показан вариант привлечения новых источников лексикографических данных в условиях Единого цифрового пространства знаний. Проблема неполноты лексикографических данных,

актуальных для решения задач поиска в цифровом пространстве, сегодня решается с помощью расширения поисковых образов за счёт привлечения ключевых слов, словарей, онтологий, таксономий, терминологических систем и синонимических рядов. Эти ресурсы и составляют основу лексикографического многообразия данных для информационного поиска и привлекаются на основе тезаурусов и их правил.

Если уж поднимать вопрос о семантически значимом представлении научных знаний, извлекаемых из цифровой среды, то опираться нужно на чёткую структуру организации научных данных. Такая организация существует в иерархических структурах классификаций. Эти структуры пронизаны горизонтальными связями иерархических уровней, через которые и возможна их верификация. Реализация таких связей в верифицированных иерархических структурах делает возможным поиск информации по классификационным индексам.

РЕСУРСЫ ЛЕКСИКОГРАФИЧЕСКОГО МНОГООБРАЗИЯ

Основными источниками для получения дополнительных лексикографических данных являются: классификационные системы, таксономии, рубрикаторы, словари предметных рубрик, онтологии, терминологические системы, авторитетные файлы имён (АФИ), словари синонимических рядов, которые должны дополнять лексикографию соответствующих семантически тождественных классов.

Классификационная система определяется в стандартах как иерархическая система классов или категорий, предназначенная для систематизации информационных ресурсов любых видов. В классификационной системе каждому понятию присваивается уникальное обозначение. Системы подразделяются на перечислительные и синтетические (аналитико-

синтетические). Последние отличаются от первых тем, что в них предусмотрена возможность синтезировать составные классификационные коды пользователем. Фасетные классификационные системы, по нашему мнению, следует рассматривать как самостоятельные классификационные схемы, которые также являются синтетическими, поскольку в них индексы сложных тем синтезируются путём объединения простых индексов. Отдельные библиографические классификации сочетают иерархический и фасетный принципы организации классов, хотя, как показывает современный опыт, это не самый удачный путь их совершенствования. Примером является структура Универсальной десятичной классификации (УДК), описанная в стандарте ГОСТ 7.90-2006 [8], которая сегодня содержит количество индексов, организованных по фасетному принципу, равное количеству индексов основного ряда этой классификации. Иерархические структуры тезаурусов и классификационных систем, как правило, не совпадают, поскольку в тезаурусах иерархия отражает преимущественно родовидовые и партитивные отношения, в то время как задача библиографических классификаций – отражать соотношение совокупностей документов при их распределении по иерархическим уровням или по категориям релевантности [11].

Таксономиями называются классификационные системы, используемые для организации контента чаще всего в среде Интернета. Этот термин используется для обозначения достаточно большого класса классификационных схем, которые в отечественной информационной практике называются *рубрикаторами* или *рубрикациями*. К этому же типу следует относить результаты такого процесса, как категоризация, получившего сегодня широкое распространение в качестве замены иерархическим классификациям. Среди множества таксономий находятся такие упрощённые структуры, как рубрикации тематических категорий библиографических баз данных Scopus, Web of Science, а также Государственный рубрикатор научно-технической информации (ГРНТИ).

Словари предметных рубрик – это системы, предназначенные для представления тем (предметов), обсуждаемых в документах. Они имеют форму перечисления (обычно алфавитного) заголовков тем с их подразделениями на 1 – 3 уровня и, возможно, указаниями на связи между рубриками и правилами, когда и как можно объединять разные рубрики при индексировании документов. Основная функция словарей предметных рубрик – систематизация коллекций документов на основе их содержания.

Онтологии обычно применяются как формальное описание предметной области для уточнения определённых понятий и/или определения общей терминологической базы этой предметной области. Они являются системами для представления логически выводимых утверждений с использованием предикатов первого порядка. Концепция онтологий возникла в информатике вне интересов информационного поиска. Как следствие, их структура и свойства изначально не предназначались для поиска информации, а были информационным компонентом данных для систем искусственного интеллекта. Вместе с тем, по-

скольку информационный поиск входит в число задач искусственного интеллекта, онтологии дают возможность трансформации своих свойств под нужды информационного поиска. Поэтому они могут рассматриваться как обобщение словарей информационно-поисковых языков с включением дополнительных средств отображения предметной области научного знания. Основной элемент онтологии – это конструкт (класс) с набором ограничений на свойства членов класса (индивидов). Классы определяются свойствами своих членов, имеющими характер либо атрибутов, либо отношений к другим классам. В онтологиях используются также наборы аксиом. Аксиомами могут быть утверждения о существовании класса или о наличии у класса определённых свойств, или утверждения, применяемые к отдельным индивидам. Онтологии всегда ориентированы на конкретную предметную область и не могут строиться на основе абстрактных представлений. Необходимо также сознавать, что не существует такой предметной области, для которой можно было бы создать единственно правильную онтологию.

Терминосистема определяется как набор обозначений, принадлежащих одному специальному языку. Терминосистемы не служат для поиска информации. Система управления терминологическими данными описывается в стандарте ISO 30042:2008 [12], в котором выделяются два класса понятий – «единичное понятие» (соответствует уникальному объекту) и «общее понятие» (которое соответствует классам объектов). У понятия есть определение, основная функция которого – объяснение его для пользователя; и в этом смысле определение можно сопоставить с лексическими примечаниями тезауруса. Для идентификации понятия используются обозначения. Если два термина являются синонимами, то один из них обычно следует установить как «предпочтительный» (ГОСТ Р ИСО 704–2010) [13].

Авторитетные файлы имен (АФИ) – это наборы имён объектов с целью их уникальной идентификации и предоставления доступа к их вариантным написаниям. Имя может иметь несколько вариантов, и среди них обычно выбирается одно предпочтительное. В информационном поиске АФИ могут использоваться для согласованности записей в информационных системах, просмотра и выбора поисковых терминов, для автоматического извлечения именованных сущностей из текста. Тезаурусы и АФИ имеют разные цели, однако в некоторых случаях их можно сопоставить, и тогда допустимо установление отношения точной эквивалентности между именем и элементом тезауруса.

Словари синонимических рядов – это списки терминов, которые можно заменять друг на друга (по крайней мере, в некоторых контекстах). Они могут использоваться для поиска, но не для индексирования. Термины из синонимических рядов не имеют реляторов и не классифицируются как предпочтительные/альтернативные, что обусловлено целями таких словарей – они нужны в первую очередь для сопоставления синонимических терминов.

Перечисленный набор источников не является исчерпывающим и представляет ресурсы, обсуждаемые

в соответствующих ГОСТах и часто используемые в практике информационного поиска, который обычно организуется в тезаурусах на базе онтологий [14–16].

МЕТОДОЛОГИЧЕСКАЯ МОДЕЛЬ СОПОСТАВЛЕНИЯ ЛИНГВИСТИЧЕСКИХ РЕСУРСОВ

Представленные источники расширения лексикографического арсенала не могут напрямую применяться в информационном поиске. Тому причиной – особенности организации этих ресурсов и различия в структуре связей используемых данных. Их сопоставление можно осуществить, опираясь на требования международного стандарта ISO 25964-2 [17], в котором описаны принципы сопоставления структур и элементов словарей информационно-поисковых языков для поиска в пространстве разнородных информационных ресурсов с разными средствами доступа к тезаурусам. Структура связей между сопоставляемыми источниками может иметь три модели.

Модель единой структуры требует полного взаимно однозначного соответствия всех элементов каждого словаря элементам других сопоставляемых словарей. Такое соответствие обеспечивает эквивалентность поисковых процессов во всех информационных ресурсах, обслуживаемых словарями единой структуры. Однако полное соответствие возможно лишь при централизованной разработке всех словарей одновременно, либо при использовании одного из них в качестве жёсткой модели для разработки других словарей. В таких случаях мы фактически имеем единый словарь понятий, выраженных в разных словарях разными языковыми средствами. Модель единой структуры рекомендуется именно для многоязычных тезаурусов и классификаций. При сопоставлении словарей, разработанных независимо друг от друга для информационных ресурсов с различным содержанием, добиться единой структуры невозможно, даже при внесении в сопоставляемые словари существенных изменений.

Модель прямых соответствий устанавливает соответствие элементов каждого словаря с элементами всех других сопоставляемых словарей, структуры которых не идентичны друг другу и могут иметь значительные различия в типах и объёмах. В такой модели каждый словарь может использоваться для поиска по ресурсам, обслуживаемым всеми другими словарями. Для этого элементы каждого словаря должны иметь прямое соответствие во всех других словарях. Применение модели прямых соответствий предпочтительно, если требуется согласовать словари нескольких независимо разработанных систем. Когда словарей три и более, сеть связей отдельных словарных единиц становится необозримой и требует специализированной компьютерной поддержки.

Модель звёздной структуры – здесь один из сопоставляемых словарей выступает как центральный, и с ним сопоставляются понятия других словарей. В такой модели каждое понятие в центральном словаре должно быть сопоставлено с понятиями словарей-спутников и наоборот. Это обеспечивает поиск по центральному словарю во всех ресурсах словарей-спутников, которые свою очередь обеспечивают по-

иск в ресурсе центрального словаря. Следует заметить, что когда связи между словарями-спутниками нарушаются, а доминирует только один словарь, поиск по периферийным словарям в ресурсе каждого из них потребует двойного перекодирования запроса: сначала к элементам центрального словаря, а затем – к элементам словаря целевого ресурса поиска. Звёздная модель рекомендуется для согласования большого числа словарей, поскольку она сводит это согласование к парным сопоставлениям.

К сожалению, установление и поддержание соответствия всех трёх моделей сопряжено со значительными трудностями, поэтому на практике часто приходится ограничиваться *частичными соответствиями словарей*, когда не все элементы каждого словаря находят себе соответствие в других источниках.

О ШКАЛАХ СООТВЕТСТВИЯ ПРИ СОПОСТАВЛЕНИИ И ГРАНИЦАХ ПРИМЕНИМОСТИ

Путь лингвистического информационного поиска оказывается чрезвычайно трудоёмким и затратным. Выявленные соответствия далеко не всегда исчерпывающие, это поднимает вопрос о необходимости шкалы для установления меры соответствия понятий. При выборе такой шкалы правильнее всего использовать те же оценки доверия, которые применяются внутри словарей и тезаурусов. Этот же путь намечен в стандарте ISO 25964-2 [17], который рекомендует при сопоставлении словарей устанавливать между их единицами те же соответствия, что и внутри словарей. Для тезаурусов наиболее типичными отношениями являются *эквивалентность*, *иерархия* и *ассоциация*. Эти же отношения применяются и к другим типам словарей. Наиболее полезным отношением для согласования словарей информационного поиска является отношение эквивалентности, которое рекомендуется устанавливать в первую очередь.

Отношение эквивалентности устанавливается между терминами двух словарей в том случае, когда обозначаемые ими понятия совпадают друг с другом. В простейшем случае в сопоставляемых словарях совпадающие понятия выражаются одинаковыми терминами. Но совпадение терминов не является достаточным условием для совпадения понятий. Например, термин «операция» в словарях по медицине, компьютерам и военному делу обозначают три разных понятия. А термины «воспаление лёгких» и «пневмония» из разных словарей, скорее всего, будут эквивалентны как обозначающие одну и ту же болезнь.

Для сопоставления словарей с различающимися структурами понятий стандарт вводит понятие *сложная эквивалентность*. Она устанавливается в том случае, когда нужно отразить соответствие термина из одного словаря понятию, выражаемому комбинацией терминов другого словаря. По характеру отношений между объёмами понятий сложную эквивалентность можно подразделить на *эквивалентность пересечения* (в случае, когда объём понятия исходного термина выражается пересечением – общей частью понятий комбинируемых терминов другого словаря) и *эквивалентность объединения* (если сложное понятие соответствует сумме комбинируемых понятий).

Кроме того, в стандарте вводятся понятия точной и неточной эквивалентности. *Точная эквивалентность* имеет место тогда, когда понятия взаимозаменяемы во всех приложениях, допустимых для понятий. *Неточная эквивалентность* бывает в тех случаях, когда понятия взаимозаменяемы только в некоторых контекстах, либо они имеют только частично пересекающиеся лексические области (например, «капитан корабля» и «шкипер»).

Иерархическое отношение устанавливается в случае, если одно из понятий (вышестоящее) шире другого (нижестоящего) по охвату темы. Основным видом иерархии в словарях информационных систем является *родовидовое отношение* между понятиями одной онтологической категории, когда объём одного понятия включает в себя объём другого (соподчинённость). Частным случаем иерархического отношения является *партиципное отношение* (часть – целое). *Родственное отношение* – связь между множеством и элементом множества – во многих словарях также рассматривается как вид иерархии. Иерархические отношения являются наиболее строгими и однозначными, поскольку соподчинённость понятий – это основополагающее требование иерархии.

Наконец, *ассоциативные отношения* устанавливаются между понятиями в случаях, когда между ними нет прямого эквивалентного или иерархического соответствия, но они явно связаны в предметной области. Такое отношение может быть полезно для индексирования сходных по тематике документов в задачах расширенного поиска информации.

Следует учитывать, что поиск соответствия по классификациям возможен только там, где главные классификационные принципы совпадают. Примеров рубрикаторов и классификаций, различающихся в принципах построения, много. Например, классификационные схемы компании *Clarivate*, утверждающей, что её классификации строятся индуктивно, «снизу вверх», в отличие от якобы устаревших классификаций, использовавших дедуктивную логику «сверху вниз». Другим примером может служить подход Организации экономического сотрудничества и развития (ОЭСР) в разработанном «Руководстве Фраскати» (1963 г., новая редакция 2007 г.). Эта классификация отличается от академических классификационных схем, поскольку категории, составляющие «Руководство», выделяются на основе их прикладной функциональности, а не тематического содержания. Очевидно, что при сопоставлении таких ресурсов, при сходстве данных, мы получим противоречивые результаты [15].

Алгоритмически организация лингвистического поиска с расширением словарно-фразеологической базы сложна для воплощения, однако она безусловно необходима. Эта необходимость нарастает в связи с продолжающимися разработками по семантическому Web и углублением исследований по полнотекстовому поиску. Шкалы соответствия, предназначенные по большому счёту для борьбы с полисемией, предоставляют линейку для сравнения понятий и исключения семантических разрывов в тематических описаниях.

ПУТЬ СЕМАНТИЧЕСКОГО ИНФОРМАЦИОННОГО ПОИСКА

Семантический информационный поиск опирается на реализацию дорожной карты вида: *описание класса – индекс класса – библиографическая запись* (см. рис. 2). Это означает, что мы используем имеющийся в распоряжении лингвистический и лексикографический арсенал для указания релевантной темы, далее определяем её индекс и применяем этот индекс как указатель для поиска индексированного объекта. В качестве поискового образа в случае семантического поиска выступает индекс класса, а расширение поискового образа достигается привлечением данных таблиц сопоставления [18]. Семантический информационный поиск, безусловно, повышает точность нахождения объекта, а релевантность выборки очевидна и не требует обсуждения.

Семантический поиск, опирающийся на таблицы сопоставления классификационных индексов, обеспечивает устойчивую связь классификаций и позволяет сохранять целостность диады *описание класса – индекс класса*. Он по своей сути представляет собой поиск по классификационным индексам на основании данных таблиц сопоставления действующих классификационных систем [19].

Таблицы сопоставления – это результат сравнения индексированных понятий из разных наиболее авторитетных классификаций. Процесс сопоставления через индексы различных классификационных систем (УДК, ББК, ДКД, ГРНТИ, Рубрикатор ВИНТИ) позволяет выражать одно конкретное понятие и указывает общие для этих классификаций «реперные» точки в индексированном пространстве Универсума знаний.

Результат такого сопоставления представляет собой группу классификационных индексов, обозначающих семантически тождественные понятия анализируемых универсумов. Использование этих индексов как указателей на объект поиска не только избавляет от полисемии, но и лишает информационное пространство языковых барьеров, позволяя абстрагироваться от лексикографической нагрузки.

Результаты сопоставления также можно использовать как ресурс для расширения лексикографического арсенала поискового образа, поскольку установленное семантическое тождество индексов позволяет предполагать семантическое тождество описаний, соответствующих этим индексам. В этом случае конъюнкция описаний расширяет лексикографию поискового образа, которая становится многообразнее, поскольку появляется возможность расширять поисковый лексикографический потенциал описаниями тождественных индексов иных классификаций [20]. Таким образом, семантический путь информационного поиска также может служить расширению лексикографической основы поискового образа через объединение классификаций-спутников.

ВЫВОДЫ

1. Необходимость расширения лексикографического потенциала поискового образа документа диктуется растущим многообразием лексики, использу-

емой в описаниях объектов Единого цифрового пространства научных знаний, и отвечает принципу необходимого разнообразия (закон Эшби).

2. Успешность информационного поиска в Едином цифровом пространстве научных знаний достигается сочетанием приёмов тематического и семантического поиска в зависимости от задач, стоящих перед исследователем.

3. Единое цифровое пространство в части описания документов и объектов необходимо унифицировать, опираясь на опыт библиотек, хранилищ данных, обширных электронных каталогов и стандартов международного уровня. Его должны составлять документы, которые индексированы с использованием действующих классификаций.

СПИСОК ЛИТЕРАТУРЫ

1. Антопольский А.Б., Калёнов Н.Е., Серебряков В.А., Сотников А.Н. О едином цифровом пространстве научных знаний // Вестник Российской Академии Наук. – 2019. – Т. 89, № 7. – С. 728-735.
2. Атаева О.М., Калёнов Н.Е., Серебряков В.А. Онтологический подход к описанию единого цифрового пространства научных знаний // Электронные библиотеки. – 2021. – Т. 24, № 1. – С. 3-18.
3. Скарук Г.А. Языки тематического поиска в электронном каталоге и индивидуальное поисковое взаимодействие // БИБЛИОСФЕРА. – 2005. – № 2. – С. 45-49.
4. Сукиасян Э.Р. Классификационная система как инструмент автоматизированного поиска. Часть 1. Почему классификационный поиск эффективнее вербального? // Научные и технические библиотеки. – 2015. – № 3. – С. 20-29.
5. Deleuze G., Guattari F. Rhizome. Introduction. P., 1976. – P. 17-38.
6. ГОСТ 7.49-84 Рубрикатор ГАСНТИ. Структура, правила использования и ведения / В. Н. Белоозеров, В. А. Глинский, Е. Н. Казаков [и др.]; Государственный комитет СССР по стандартам. – Москва: Изд-во стандартов, 1984. – 22 с. (Система стандартов по информации, библиотечному и издательскому делу).
7. ГОСТ Р 7.0.49–2007 Государственный рубрикатор научно-технической информации. Структура, правила использования и ведения / Федеральное агентство по техническому регулированию и метрологии. – Переиздание. – Москва: Стандартинформ, 2018 (Система стандартов по информации, библиотечному и издательскому делу).
8. ГОСТ 7.90–2006 Универсальная десятичная классификация. Структура, правила использования и индексирования / МТК 191. – Введён 2007-07-01 (Система стандартов по информации, библиотечному и издательскому делу).
9. ГОСТ 7.25–2001 Тезаурус информационно-поисковый одноязычный. Правила разработки, структура, состав и форма представления / МТК 191. – Введён 2002-07-01. – 16 с. (Система стандартов по информации, библиотечному и издательскому делу).
10. ГОСТ Р ИСО 15489-1–2019 Информация и документация. Управление документами. Часть 1. Понятия и принципы: национальный стандарт Российской Федерации / Федеральное агентство по техническому регулированию и метрологии. – Дата введения 2020-01-01 (Система стандартов по информации, библиотечному и издательскому делу).
11. Белоозеров В.Н., Антошкова О.А. Разработка национального стандарта на сопоставление тезаурусов с другими классификационными языками // Научное издание международного уровня – 2019: Стратегия и тактика управления и развития: Материалы 8-й Междунар. научно-практич. конф. (Москва 23 – 24 апреля 2019 г.). – Екатеринбург, 2019. – С. 35-43.
12. ГОСТ Р ИСО 30042-2016 (ISO 30042:2008) Системы управления терминологией, базами знаний и контентом. Обмен терминологическими базами [TermBase eXchange (TBX)] Systems to manage terminology, knowledge and content. TermBase eXchange (TBX). Москва: Стандартинформ, 2020.
13. ГОСТ Р ИСО 704-2010 ТЕРМИНОЛОГИЧЕСКАЯ РАБОТА. Принципы и методы / ISO 704:2009 Terminology work — Principles and methods (IDT). – Издание официальное. – Москва: Стандартинформ. – 2012.
14. Гусенков А.М., Бухараев Н.Р., Биряльцев Е.В. Построение онтологии предметной области для корпоративного веб-приложения // Научный сервис в сети Интернет: труды XXI Всероссийской научной конференции (23-28 сентября 2019 г., г. Новороссийск). — Москва: ИПМ им. М.В. Келдыша, 2019. — С. 302-317. — URL: <http://keldysh.ru/abrau/2019/theses/58.pdf> doi:10.20948/abrau-2019-58
15. Медведева И.Б. Построение онтологических моделей систем автоматизированного проектирования // Информатика, вычислительная техника и инженерное образование. – 2015. – № 3(23). – С. 1-7.
16. Шебалов Р., Горшков С. Современные российские разработки в области онтологического моделирования. Обзор / © ООО «ТриниДата», 2018, 30 с.
17. ISO 25964-2:2013 Information and documentation — Thesauri and interoperability with other vocabularies — Part 2: Interoperability with other vocabularies : International Standard / ISO. – Geneve, 2013. – 99 с.
18. Baker T., et al. W3C Library Linked Data Incubator Group Final Report. W3C, 25 October 2011. – URL: <http://www.w3.org/2005/Incubator/ld/XGR-ld/>
19. Скарук Г.А. Возможности поиска по классификационным индексам в электронном каталоге // Научные и технические библиотеки. – 2016. – № 3. – С. 19-29.
20. Шомшор М., Адамс Д., Пиндлбери Д., Роджерс Г. Отчет о международном исследовании. Классификация данных: как делать осознанно

знанный выбор, ведущий к желаемым результатам / Мартин Шомшор (Martin Szomszor), Джонатан Адамс (Jonathan Adams), Дэвид Пендлбери (David A. Pendlebury) и Гордон Роджерс (Gordon Rogers), Институт научной информации (ISI), 2021. – 22 с. – URL: webofsciencegroup.com/isi

21. Антопольский А.Б., Белоозеров В.Н., Макарова Т.С., Дмитриева Е.Ю. Установление соответствий рубрик ГРНТИ рубрикам других систем классификации // Научно-техническая информация. Сер. 1. – 2015. – № 3. – С. 23-26.

Материал поступил в редакцию 04.02.22.

Сведения об авторах

МЕТЛОВА Ася Владимировна – главный специалист Научно-методологического отделения ВИНТИ РАН, Москва
e-mail: asjavladimirovna@gmail.com

АСТАХОВА Татьяна Сергеевна – старший научный сотрудник НМО ВИНТИ РАН
e-mail: ran-astahova@mail.ru

СЫСОЕВ Александр Николаевич – главный специалист НМО ВИНТИ РАН
e-mail: vims-sysoev@mail.ru