

Оригинальная статья / Original paper

<https://doi.org/10.24069/SEP-25-34>



Современное состояние нормативной базы поискового индексирования информационных ресурсов издательств и библиотек

И. В. Тимошенко  

Московский государственный институт культуры, г. Москва, Российская Федерация

 timigor@gpntb.ru

Резюме. В статье анализируются ключевые принципы разработки поисковых тезаурусов и методы индексирования, используемые в библиотечно-издательской сфере. Основные концепции построения информационно-поисковых тезаурусов – специализированных словарей, применяемых для индексирования ресурсов в словарных информационно-поисковых системах, основаны на использовании метаданных – поисковых индексов для определения тематической релевантности между проиндексированными информационными ресурсами и поисковыми запросами. В области информационного поиска в последние годы большое распространение получили информационно-поисковые языки ключевых слов, основанные на специализированных словарях с контролируемой лексикой – информационно-поисковых тезаурусах. Рассматриваются основные этапы развития тезаурусов – от построения структуры поисковых терминов и выражаемых ими понятий до формирования систематизированной структуры знаний, выражаемых терминами. Показано, как отражается развитие концепций построения информационно-поисковых тезаурусов и поискового индексирования в нормативной базе организации знаний, принятой в сети Интернет и развиваемой Консорциумом Всемирной паутины (W3C), а также на уровне организаций международной (ISO) и российской национальной (ГОСТ) стандартизации. Сделан вывод о необходимости обновления российской национальной нормативной базы поискового индексирования путем гармонизации российских и международных стандартов в этой области.

Ключевые слова: информационные ресурсы, информационный поиск, тезаурусы, поисковое индексирование, стандартизация, нормативная база, Интернет

Для цитирования: Тимошенко И. В. Современное состояние нормативной базы поискового индексирования информационных ресурсов издательств и библиотек. *Научный редактор и издатель.* 2025;10(1):154–168. <https://doi.org/10.24069/SEP-25-34>

Indexing of publishers and libraries information resources: the current state of the regulatory framework

I. V. Timoshenko  

Moscow State Institute of Culture, Moscow, Russian Federation

 timigor@gpntb.ru

Abstract. The author analyzes the key principles of developing search thesauri and indexing methods used in the library and publishing industry. The basic concepts of building information retrieval thesauri, specialized dictionaries used for indexing resources in dictionary information retrieval systems, are based on the use of metadata, search indexes, to determine the thematic relevance between indexed information resources and search queries. The author shows that in the field of information search in recent years,

keyword search languages based on specialized dictionaries with controlled vocabulary, information retrieval thesauri, have become widespread. The author examines the main stages of the development of thesauri – from the construction of the structure of search terms and the concepts expressed by them, to the formation of a systematic structure of knowledge expressed by terms. The author discusses the reflection of the development of concepts for building information retrieval thesauri and search indexing in the regulatory knowledge organization framework adopted on the Internet and developed by the World Wide Web Consortium (W3C), as well as at the level of international (ISO) and Russian national (GOST) standardization organizations. The author concludes that it is necessary to update the Russian national regulatory framework for search indexing by harmonizing Russian and international standards in this area.

Keywords: thesaurus, indexing, information resources, information retrieval, standardization, regulatory framework, Internet

For citation: Timoshenko I. V. Indexing of publishers and libraries information resources: the current state of the regulatory framework. *Science Editor and Publisher*. 2025;10(1):154–168. <https://doi.org/10.24069/SEP-25-34>

ВВЕДЕНИЕ

Развитие цифровых технологий и переход к сетевым формам хранения и распространения информации привели к радикальному пересмотру традиционных подходов к организации знаний и индексированию информационных ресурсов. В условиях экспоненциального роста объемов электронных данных и повсеместного внедрения автоматизированных информационно-поисковых систем (ИПС) возрастает значимость лингвистических средств, обеспечивающих точную идентификацию и релевантный доступ к содержанию документов. Особенно остро этот вызов стоит перед библиотечно-издательским сектором, где эффективность поиска и навигации напрямую зависит от качества использованных средств индексирования, в том числе от уровня развития и согласованности применяемых тезаурусов и классификаций.

Ключевым инструментом, обеспечивающим когерентность терминологии и повышение семантической точности, остаются информационно-поисковые тезаурусы (ИПТ), которые исторически формировались в рамках национальных и международных стандартов. Однако в последние два десятилетия произошел качественный сдвиг: традиционные модели построения тезаурусов, основанные на ручном управлении терминологией и печатной парадигме документа, уступают место онтологическим моделям, совместимым с концепцией Семантической паутины [1]. Международное сообщество отреагировало на эти изменения созданием новых стандартов – в первую очередь ISO 25964 и SKOS (Simple Knowledge Organization System), в которых термин уже не отождествляется с понятием, а репрезентация знаний опирается на формализованные структуры RDF-графов.

На этом фоне российская нормативная база, сформированная преимущественно в 1980–1990-е гг., сохраняет приверженность концептуально устаревшим положениям стандартов ISO 2788 и ISO 5964, отмененных на международном уровне еще в 2011 г. Такие стандарты, как ГОСТ 7.24–80, ГОСТ 7.25–80, ГОСТ 7.59–2003 и ГОСТ Р 7.0.66–2010, хотя и сыграли важную роль в формировании современных подходов к библиотечному индексированию, в настоящее время не отражают ни структурных изменений в коммуникационной среде, ни технологических возможностей современных цифровых платформ [2; 3].

Таким образом, настоящая статья исходит из позиции, согласно которой дальнейшее развитие отечественной нормативной базы поискового индексирования невозможно без ее системного пересмотра в направлении гармонизации с международными моделями, прежде всего ISO 25964 и SKOS. Автор выступает с экспертной позицией, основанной на участии в работе технического комитета ISO/TC 46, и формулирует практические рекомендации, направленные на интеграцию российских стандартов в глобальное пространство систем организации знаний.

ОСНОВНЫЕ ПРИНЦИПЫ ПОИСКОВОГО ИНДЕКСИРОВАНИЯ ИНФОРМАЦИОННЫХ РЕСУРСОВ

Исторически сложились два метода поискового индексирования при формировании библиографических записей в электронных каталогах в зависимости от типа используемых информационно-поисковых языков (ИПЯ): предкоординатный и координатный (посткоординатный) [4–6].

Предкоординатный метод предполагает предварительную разработку некоторой схемы знаний,

где знания представляют собой определенные понятия, выраженные как отдельными терминами, так и грамматическими конструкциями. Такие схемы, как правило, имеют иерархическую организацию, структурированную по тематическим направлениям. В практике библиотечной и издательской деятельности исторически выделяют два основных вида таких схем: тематический рубрикатор и классификационную систему, которые различают по форме представления тематических рубрик. Для рубрикаторов вербальные наименования рубрик эквивалентны их классификационным кодам, рубрики классификационных систем дополнительно кодируются. К основным видам классификационных систем относятся:

1) *перечислительные* – список классов, не связанных друг с другом отношениями подчинения. Такие классификации применяются для простой структуризации коллекций в виде последовательности непересекающихся видовых множеств;

2) *моноиерархические* – иерархический граф тематических рубрик с одной вершиной. Все рубрики строго определяют тематические области и связаны отношением подчинения только с одной вышестоящей рубрикой;

3) *политематические (фасетные)* – ограниченное множество моноиерархических классификаций (аспектов), построенных по независимым видовым признакам. Такие классификации представлены в развитых классификационных системах, применяемых в издательской и библиотечной деятельности. Примерами являются Универсальная десятичная классификация (УДК) и Библиотечно-библиографическая классификация (ББК).

Классификационные коды используются для формирования поисковых индексов, которые присваиваются информационным ресурсам. Каждый код уникально соответствует разделу знаний в рамках применяемой схемы и формируется исходя из анализа и оценки информационного ресурса. В дальнейшем информационный поиск осуществляется по схеме знаний без привлечения проиндексированных ресурсов.

Предкоординатный метод индексирования сформировался в докомпьютерную эпоху, начиная с конца XIX в., и был ориентирован на применение в системах карточных каталогов, представляющих собой разновидность механических поисковых систем. В электронной среде предкоординатный метод использовался в первых ИПС Интернета – поисковых каталогах. В настоящее время поисковые каталоги представлены преимущественно в поисковых системах локальных

коллекций издательств, библиотек и архивов в качестве вспомогательных средств поиска. Формирование и присвоение кодированных индексов в таких системах производится вручную, возможная точность поиска зависит от правильности их присвоения, что при условии тематической многоаспектности информационных ресурсов представляет собой нетривиальную задачу и требует от индексатора определенной квалификации и навыков кодирования.

Координатный (посткоординатный) метод поискового индексирования основан на ИПЯ ключевых слов и используется при полнотекстовом поиске информационного ресурса. Полнотекстовый поиск предполагает анализ самого ресурса для оценки его соответствия информационному запросу. Такой подход отличает не только принципиальная возможность достижения абсолютной точности в определении соответствия ресурса информационной потребности пользователя, но также и значительно большая трудоемкость, что не позволяет его широко применять в сколько-нибудь обширных коллекциях информационных ресурсов. В силу большой трудоемкости такого поиска его практическое применение обычно связано с предварительной обработкой текста, в ходе которой выявляются элементы, компактно отражающие его тематику и смысловое содержание. Традиционные формы такой обработки: реферирование, аннотирование и выделение ключевых слов. В дальнейшем индексирование ресурса ключевыми словами составляет сущность координатного индексирования, а информационный поиск осуществляется путем анализа поисковых индексов без привлечения самих ресурсов, так же как при предкоординатном виде поиска.

Методы координатного индексирования получили широкое распространение начиная с 1960-х гг., когда началось активное внедрение ЭВМ во все виды деятельности, включая информационные. Карточные поисковые каталоги начали замещаться электронными, ручные методы обработки информации – автоматизированными с использованием компьютерной обработки. Моделирование процессов организации и поиска больших массивов информации в электронной среде привело к созданию теории информационного поиска [7], развивающей аналитические модели поиска, основанные на применении информационных запросов, выраженных на естественных языках и обрабатываемых компьютерными методами. Задача сопоставления ключевых слов, выделяемых из поисковых запросов на естественном языке, с ключевыми словами, которы-

ми проиндексированы ресурсы ИПС, на уровне концепции просто решается четкими алгоритмическими методами, реализуемыми на компьютерах сравнительно небольшой вычислительной мощности. Ее практическая реализация была доступна уже для ЭВМ поколений 1960–1970х гг.

Вместе с тем простая концепция поиска по ключевым словам сталкивается с проблемой вариативности, которая выражается в избыточности естественных языков – в проявлении словарной синонимии и гипонимии. Для сокращения разнообразия ключевых слов, выражающих основную тематику информационного ресурса, широко применялись дескрипторные словари, получившие дальнейшее развитие как ИПТ, в которых группы лексических единиц – терминов, связанных семантическими отношениями (синонимы, гипонимы), – организованы в иерархическую структуру. Сокращение разнообразия лексики информационных ресурсов и поисковых запросов достигается за счет ее формализации путем замены лексических единиц, включенных в ИПТ и выделенных как ключевые слова, обобщающими терминами – дескрипторами, однозначно обозначающими группу близких по значению лексических единиц [8; 9]. Для создания ИПТ необходимо определить терминосистему конкретной предметной области с последующим подбором и группировкой вокруг установленных терминов их синонимов, извлекаемых из корпуса текстов, относящихся к этой области, – сформировать иерархическую структуру *классов условной эквивалентности*. Следует отметить, что структура такого тезауруса с учетом возможных иерархических связей, определяемых отношениями «ВТ – выше», «НТ – ниже», может быть представлена как обобщенная иерархия [2]:

$$H = \{h_j\},$$

где j – идентификатор элемента иерархии H ,

$$j = (k_i)_{i=1}^{n_j},$$

здесь n_j – уровень иерархии элемента h_j ; k_i – идентификатор элемента на h_j -м подуровне иерархии H_j .

Представленная аналитическая модель аналогично описывает структуры традиционных тезаурусов, а также классификационных рубрикаторов или основных таблиц универсальных классификаций, которые можно отнести к *формальным таксономиям*. Это обстоятельство сегодня стирает принципиальные различия между двумя подходами к формированию поисковых индексов, ранее традиционно противопоставляемыми друг

другу: предкоординатным и координатным индексированием. Обобщающим понятием, объединяющим оба вида индексирования, является *онтология* как формальная спецификация концептуальной модели некоторой предметной области. В таком представлении классификации и тезаурусы можно рассматривать как сходные виды онтологий [8, с. 104–105].

ПЕРВЫЕ СТАНДАРТЫ ИПТ

В 1960-е гг. в связи с широким внедрением ЭВМ в информационные сферы деятельности и развитием словарных методов информационного поиска ИПТ получили широкое распространение в мире. Наиболее известным ИПТ, появившимся в этот период, является *Thesaurus of Engineering and Scientific Terms*, TEST («Тезаурус инженерных и научных терминов»). Впервые опубликованный в 1964 г. и переизданный в 1967 г.¹ TEST был разработан Объединенным советом инженеров при участии Министерства обороны США. Он содержит более 17 000 дескрипторов, представленных в алфавитном порядке, которые объединяют более 150 000 терминов, с разъяснениями, как они могут быть использованы для поиска информации. Важной особенностью TEST является то, что в него был включен раздел с установленными правилами формирования его структуры, которые послужили основой для дальнейших работ по стандартизации принципов построения ИПТ.

В 1972 г. в Центральном институте военно-технической информации Министерства обороны СССР (ЦИВТИ МО) был разработан и опубликован аналогичный русскоязычный «Тезаурус научно-технических терминов» под общей редакцией Ю. И. Шемакина². Отбор терминов при разработке тезауруса производился путем анализа научных и политических отечественных и зарубежных источников, в том числе справочных изданий, энциклопедий, таблиц УДК и ББК. В частности, в тезаурус был включен ряд терминов из TEST в русском переводе. Систематический указатель дескрипторов имеет трехуровневую иерархическую структуру: области, группы и дескрипторы. Всего в структуру тезауруса входили 33 дескрипторные области, включающие 301 группу из

¹ Thesaurus of Engineering and Scientific Terms. Engineers Joint Council, United States. Department of Defense. New York, NY: Engineers Joint Council; 1967. 690 p. Available from: https://books.google.ru/books?id=BHbVAAAAMAAJ&hl=ru&source=gs_book_other_versions (accessed: 25.04.2025).

² Шемакин Ю.И., ред. Тезаурус научно-технических терминов. М.: Воениздат МО СССР; 1972. 671 с. Режим доступа: <https://reallib.org/reader?file=506495&pg=3> (дата обращения: 25.04.2025).

~14 000 дескрипторов, объединяющих ~40 000 терминов. Хотя в тезаурус вошли преимущественно термины и дескрипторы военно-технической тематики, такие как «артиллерия», «военные науки», «тыл» и «снабжение» и т.д., в него также вошли общенаучные и общетехнические терминологические разделы по направлениям биологии, математики, строительства, физики, химии и т.д., что позволяет его использовать широкому кругу научных и редакционных работников.

К концу 1960-х гг. в мире и в СССР существовало уже несколько сот автоматизированных ИПС, использующих различные тезаурусы в качестве ИПЯ координатного индексирования. С учетом важности разработки единой нормативной базы для ИПТ на основании опыта, полученного рабочей группой экспертов технического комитета ИСО/ТК 46, в 1971 г. был разработан первый международный документ стандартизации – *ISO/ТК 46/Т «Recommendation 1041E of August 3, 1971»* (ИСО/ТК 46/Т «Рекомендации 1041Е»), устанавливающий принципы создания и применения тезаурусов на международном уровне. Этот документ послужил основой для ряда национальных документов стандартизации, в том числе в СССР, а затем – в Российской Федерации.

В СССР в период 1965–1973 гг. велась активная работа по систематизации существующих в стране ИПЯ, разработке и практической проверке методов унификации и максимально автоматизированного анализа основных тезаурусов, используемых в Государственной автоматизированной системе научно-технической информации (ГАСНТИ), для их согласованного ведения и гармонизированного сопоставления [9]. В 1973 г. на основе полученного опыта, а также с учетом рекомендаций, установленных в документе «Рекомендации 1041», Государственным комитетом стандартов Совета Министров СССР (Госстандарт СССР) был разработан и утвержден *ГОСТ 18383–73 «Тезаурус информационно-поисковый. Общие положения. Форма представления»*. Стандарт устанавливал самые общие требования к построению ИПТ: обязательные элементы структуры, порядок расположения терминов дескрипторной статьи и уточнения в случае их омонимии или полисемии. Стандарт распространялся на тезаурусы для обработки и поиска научно-технической информации в отраслевых ИПС и был обязателен для применения на территории СССР.

Дальнейшим развитием международной нормативной базы ИПТ стала разработка международного стандарта *ISO 2788:1974 «Documentation – Guidelines for the establishment and development of*

monolingual thesauri» (ИСО 2788:1974 «Документация. Руководство по созданию и развитию одноязычных тезаурусов»), который заменил документ стандартизации «Рекомендации 1041». В новом стандарте были более детально представлены принципы и правила, установленные ранее в TEST и в замененном документе стандартизации. Появление в 1974 г. международного стандарта ISO 2788 инициировало в 1980 г. разработку Госстандартом СССР отечественного стандарта *СТ СЭВ 174–85 «Система стандартов по информации, библиотечному и издательскому делу. Тезаурус информационно-поисковый одноязычный. Правила разработки, структура, состав и форма представления»*, заменившего *ГОСТ 18383–73*, а затем – на его основе – *ГОСТ 7.25–80* с тем же названием, переизданного в 2001 г. В стандарте был установлен более широкий набор русскоязычных и символьных обозначений связей между лексическими единицами тезауруса, включающий иерархические и ассоциативные связи, аналогичные англоязычным обозначениям. Важным положением стандарта являлось то, что он определял необходимость регистрации в нашей стране тезаурусов в соответствии с порядком, установленным ГКНТ СССР, для чего в 1978 г. ВНИТИ РАН³ совместно с ВНИИКИ⁴ был разработан нормативный документ «Порядок экспертизы и регистрации информационно-поисковых тезаурусов» [10, с. 96].

Правила построения многоязычных тезаурусов были установлены в *ГОСТ 7.24–80 «Система стандартов по информации, библиотечному и издательскому делу. Тезаурус информационно-поисковый многоязычный. Правила разработки»*, вступившем в действие с начала 1982 г. Новый стандарт расширял действие правил, установленных в *ГОСТ 7.25*, на многоязычные тезаурусы. Впоследствии он переиздавался в 1990 и 2007 гг.

Международный стандарт по многоязычным тезаурусам – *ISO 5964:1985 «Documentation – Guidelines for the establishment and development of multilingual thesauri»* (ИСО 5964:1985 «Документация. Руководство по созданию и развитию многоязычных тезаурусов») – появился позднее, он также расширял область применения основных принципов, установленных в ISO 2788, на многоязычные тезаурусы. Основные положения между-

³ Всероссийский институт научной и технической информации Российской академии наук.

⁴ Всероссийский научно-исследовательский институт классификации, терминологии и информации по стандартизации и качеству (ВНИИКИ) был частью Государственного комитета РСФСР по стандартизации, метрологии и сертификации. В настоящее время является частью Федерального агентства по техническому регулированию и метрологии (Росстандарт).

народного стандарта были аналогичны положениям ГОСТ 7.24. В 1986 г. стандарт ISO 2788 был переиздан.

Эти два стандарта стали основой международной нормативной базы формирования ИПТ более чем на два десятилетия.

Принципиальным моментом, устанавливаемым правилами первых стандартов по формированию ИПТ, являлось отождествление понятия и термина, его выражающего. В последней редакции стандарта ISO 2788:1986 эта особенность явно указывается во введении: «Для практических целей “термин” и “понятие” иногда используются взаимозаменяемо». Для указания иерархических отношений между терминами ИПТ в стандартах рекомендуется использовать теги, изначально предложенные в правилах TEST: BT (англ. *broader term* – «более широкий термин»), NT (англ. *narrower term* – «более узкий термин»), RT (англ. *related term* – «родственный термин») и т.д. Стандарты определяют еще ряд тегов, дающих возможность указания иных парадигматических связей между лексическими единицами тезауруса, что дает пользователю информацию о существовании нескольких понятий, выражаемых одним термином, при этом вся структура тезауруса ориентирована на ручной выбор термина для индексирования непосредственно пользователем.

ПЕРВЫЕ СТАНДАРТЫ ПО МЕТОДАМ ПОИСКОВОГО ИНДЕКСИРОВАНИЯ

Наряду с правилами формирования ИПТ большое значение имеют методы их применения для индексирования информационных ресурсов различных видов. Впервые такие правила были определены в 1975 г. в международном стандарте ISO 999:1975 «*Documentation – Index of a publication*» (ИСО 999:1975 «Документация. Индекс издания»). В этом стандарте были представлены краткие рекомендации по формированию поисковых индексов для информационных ресурсов различных видов: книг (включая художественные произведения), периодических изданий, отчетов, патентных документов, других печатных документов, а также других ресурсов, таких как электронные документы, фильмы, звукозаписи, видеозаписи, графические материалы, карты и трехмерные объекты и т.д. Стандарт был существенно пересмотрен в 1996 г. Основные положения, приведенные в последней редакции стандарта, имеют отношение не только к ручной подготовке индексов, но и к использованию компьютерных методов. Стандарт определяет правила формирования

для различных видов индексов: предметных, авторских, именных, географических, кодовых (номерных), а также для указателей наименований. Наиболее подробно рассмотрены технические вопросы построения индексов: выбор заголовков рубрик (входных терминов), указание адресов материала, оформление перекрестных ссылок, порядок расположения рубрик.

Стандарт ISO 999 устанавливает правила формирования индексов, но не содержит рекомендаций по методам предварительного анализа ресурсов с целью их индексирования. Эти вопросы были определены в международном стандарте ISO 5963:1985 «*Documentation – Methods for examining documents, determining their subjects, and selecting indexing terms*» (ИСО 5963:1985 «Документация. Методы анализа документов, определения их тематики и выбора условий индексирования»).

В стандарте представлены общие методы анализа документов на этапах анализа и определения их тематики, которые могут быть применимы для всех видов индексирования – как терминами, определяемыми тезаурусом, так и кодами классификационных систем. При этом особо отмечается, что описанные методы применимы в ручной обработке текстовых ресурсов. Практические рекомендации, приведенные в стандарте, относятся к процессам определения целей индексирования, основных понятий, характеризующих содержание документа, анализа документа, подбора терминов индексирования и проверки их качества. В качестве графической иллюстрации дана блок-схема процесса индексирования с использованием тезауруса. Рекомендации по организации индексирования документов, приведенные в стандарте, могут быть также полезны для анализа запросов пользователей и их трансляции в контролируемые термины применяемого для индексирования ИПЯ. Описанные в стандарте методы могут быть полезны не только при определении терминов индексирования, но также при подготовке тезисов для рефератов и аннотаций [11].

В этот же период были разработаны национальные стандарты:

- 1) ГОСТ 7.44–84 «СИБИД. Систематизация документов. Общие требования»;
- 2) ГОСТ 7.45–84 «СИБИД. Предметизация документов. Общие требования».

В них были учтены положения ISO 5963, и впоследствии они стали основой для разработки межгосударственных стандартов, действовавших на территории России:

- 1) ГОСТ 7.59–90 «СИБИД. Индексирование документов. Общие требования к систематизации

и предметизации» (стандарт был пересмотрен и действует как ГОСТ 7.59–2003);

2) ГОСТ 7.66–92 «СИБИД. Индексирование документов. Общие требования к координатному индексированию».

ГОСТ 7.66 имел более широкую область действия в сравнении с международным стандартом, наряду с основными положениями ISO 5963 в нем даны рекомендации по фактографическому индексированию, а также по технологии автоматизированного индексирования. В настоящее время в России действует пересмотренная национальная версия этого стандарта (ГОСТ Р 7.0.66–2010), которая не содержит принципиальных изменений по отношению к предыдущей.

Общим для всех разработанных в этот период и действующих по настоящее время стандартов по методам индексирования было то обстоятельство, что они разрабатывались непосредственно перед широким распространением Интернета, поэтому к моменту публикации они воспринимались и воспринимаются в настоящее время как руководства по индексированию печатных изданий по правилам, применимым исключительно к ручному индексированию.

ЭВОЛЮЦИЯ КОНЦЕПЦИИ ПОСТРОЕНИЯ ТЕЗАУРУСОВ

Отождествление понятия и термина, его выражающего, принятое в первых стандартах ИПТ, не имеет большого значения в случае, когда подбор термина из тезауруса осуществляется при непосредственном участии человека. В случае обработки текстов автоматизированной информационной системой условия для интерпретации используемых терминов должны быть определены более полно.

Соответствие между понятием и термином человек определяет интуитивно, исходя из контекста предметной области и, как правило, может адекватно интерпретировать термин в результатах поиска. Исходное понятие существует только в представлении пользователя и в общем случае не зависит от его выражения термином. Термин выполняет роль кода, обозначающего исходное понятие, и необходимость в нем появляется только в процессе коммуникации между людьми, а в современных условиях – и при участии автоматизированных систем, которые при выборе термина действуют строго в рамках установленных правил и, в отличие от человека, не имеют информации о контексте.

Существующая проблема омонимии естественных языков может приводить к ошибкам

и ложным выводам при выборе терминов автоматизированной системой, что можно проиллюстрировать на примере термина «марс», который, в зависимости от контекста, может обозначать планету, шоколадный батончик или наименование общероссийского проекта ассоциации Арбикон⁵ по межрегиональной аналитической росписи статей.

Развитие современных информационных технологий создало новые условия и обусловило новые требования к лингвистическим средствам ИПС. Сегодня область применения поискового индексирования существенно расширилась и обеспечивает доступ к информации всех видов во всех форматах, включая сетевые информационные ресурсы, как локальные, так и удаленные, размещенные в Интернете. Развитие электронных средств коммуникации и глобальной сети Интернет привело к необходимости интенсификации семантической обработки информационных ресурсов разных видов, поступающих из разных источников, для обеспечения возможности их эффективного поиска. Решение этой задачи связано с преодолением двух проблем.

Во-первых, необходимо обеспечить возможность извлечения ключевых терминов из множества сетевых ресурсов не только с участием человека, но и с применением технических средств, что диктует необходимость более четкого определения понятия как объекта и формализации выбора адекватного ему термина для индексирования ресурса. Это возможно путем развития моделей представления понятия как формализованного набора данных, устанавливающего в широком контексте лексическую единицу тезауруса – термин, применяемый в конкретных предметных областях знания или деятельности [1].

Во-вторых, развитие информационных сервисов Интернета способствовало появлению большого разнообразия контролируемых словарей и видов лексики, используемых для навигации в разнообразных сетевых ресурсах, поиска и фильтрации информации в них. Наряду с потребностью семантической обработки самих информационных ресурсов, появилась потребность в определении требований функциональной совместимости между собой различных тезаурусов, а также их совместимости с другими словарями, применяемыми в различных информационных

⁵ Поиск статей в российской периодике (МАРС). Ассоциированные Региональные Библиотечные Консорциумы (АРБИКОН). Режим доступа: https://web.archive.org/web/20221004223034/https://arbicon.ru/services/mars_analitic.html/ (дата обращения: 24.06.2025).

системах Интернета. В задачах информационного поиска только реализация такой совместимости может обеспечить возможность семантической обработки всей совокупности видов информационных ресурсов, представленных в Интернете [3].

Решение указанных проблем возможно путем развития нормативной базы поискового индексирования в направлении концепции Семантической паутины, определяющей сегодня развитие информационно-коммуникационной среды Интернета. При этом именно технология индексирования преобразует неструктурированные большие данные Интернета в организованный массив, доступный для информационного поиска, что необходимо для реализации этой концепции.

Базовым понятием Семантической паутины является модель представления данных RDF (англ. *Resource Description Framework* – «структура описания ресурсов»)⁶, позволяющая описывать структуры индексированных информационных ресурсов, образующих онтологию, в виде графа, представляющего семантическую сеть. Семантические сети позволяют объединять информационные ресурсы в концептуальные модели, исходя из информационных потребностей конкретной предметной области. Примером таких моделей могут служить концептуальные модели библиографических данных [12] IFLA (англ. *International Federation of Library Associations* – Международная федерация библиотечных ассоциаций и учреждений, ИФЛА). Индексирование информационных ресурсов как компонентов онтологий знаний позволяет формировать базы знаний для получения логических выводов по правилам математической логики в задачах, реализуемых средствами интеллектуальных информационных систем.

СОВРЕМЕННОЕ СОСТОЯНИЕ НОРМАТИВНОЙ БАЗЫ ПОИСКОВОГО ИНДЕКСИРОВАНИЯ

Переход к онтологически управляемым моделям: SKOS и семантическая паутина

Необходимым условием реализации концепции Семантической паутины является переход от существующей технологии индексирования информационных ресурсов, ориентированных на традиционные «бумажные» технологии, к новой, отвечающей требованиям электронной коммуникационной среды и реализующей ее воз-

можности. Разработка нормативной базы новой технологии велась с конца 1990-х гг. Началом работ в этом направлении можно считать проекты DESIRE II, LIMBER и SWAD-Europe [13], в рамках которых разрабатывались методы представления тезаурусов на языке RDF. С 2004 г. работа была продолжена рабочей группой W3C (англ. *World Wide Web Consortium* – Консорциум Всемирной паутины). Результатом усилий стала разработка стандарта W3C – SKOS, официальный релиз которого был опубликован в 2009 г. [14].

Стандарт SKOS изначально разрабатывался как инструмент интеграции классических систем организации знаний (например, тезаурусов, используемых в библиотеках) с веб-ориентированными структурами – открытыми каталогами, блогами и другими форматами. В основе SKOS лежит концептуальная модель, где базовыми элементами выступают не термины, а абстрактные понятия (концепции), которые выражаются через эти термины. Для описания связей между лексическими единицами стандарт вводит специальные классы и свойства, характерные для традиционных тезаурусов, но адаптированные под веб-среду.

Каждая концепция в SKOS трактуется как RDF-объект (ресурс с метаданными) и может быть детализирована с помощью 32 предопределенных RDF-атрибутов. Это позволяет не только формализовать семантические отношения между понятиями (например, «род – вид» или «синонимия»), но и обеспечивает совместимость с технологиями семантического веба, упрощая интеграцию библиотечных систем с современными цифровыми платформами. Семантические отношения между понятиями, определяемые в SKOS, однозначно соответствуют иерархическим и ассоциативным отношениям, рекомендованным в ISO 2788. Как пример, три стандартных свойства: *skos:broad* («более широкое») и *skos:narrow* («более узкое») для иерархических связей и *skos:related* («связанные») для ассоциативных связей между понятиями. Концепции могут объединяться в концептуальные модели, включающие часть или все элементы контролируемого словаря относительно общих значений их свойств RDF.

При создании стандарта SKOS изначально закладывалась совместимость с нормами ISO 2788 и ISO 5964, регламентирующими базовые принципы построения тезаурусов. Однако эти документы, разработанные в эпоху доминирования бумажных носителей, не учитывают требований цифровой эпохи: особенностей электронных коммуникационных сред, специфики

⁶ W3C. Resource Description Framework (RDF) 1.1: Concepts and Abstract Syntax. W3C Recommendation. 25 February 2014. Available from: <https://www.w3.org/TR/rdf-concepts/> (accessed: 25.04.2025).

работы с данными в цифровом формате и возможностей семантической веб-среды. Это создает разрыв между устаревшими рекомендациями, ориентированными на ручное управление терминологией, и современными потребностями автоматизированной обработки информации в интернет-пространстве.

Международный стандарт ISO 25964: структура и функциональность

Появление SKOS послужило толчком к разработке нового национального стандарта Великобритании *BS 8723 «Structured Vocabularies for Information Retrieval – Guide»* (BS 8723 «Структурированные словари для информационного поиска. Руководство»)⁷, учитывающего развитие принципов индексирования ресурсов Интернета и гармонизированного со стандартами W3C. Британский стандарт стал основой для разработки международного стандарта *ISO 25964:2011 «Information and documentation – Thesauri and interoperability with other vocabularies»* (ISO 25964:2011 «Информация и документация. Тезаурусы и взаимодействие с другими словарями»), состоящего из двух частей. В состав рабочей группы вошли эксперты из 15 стран под председательством BSI (Великобритания) и с секретариатом, возглавляемым NISO (США). Со стороны России в проекте принимал активное участие представитель ТК 191 В.Н. Белоозеров (ВИНИТИ РАН). Первая часть стандарта *ISO 25964-1 «Information and documentation – Thesauri and interoperability with other vocabularies – Part 1: Thesauri for information retrieval»* (ISO 25964-1 «Информация и документация. Тезаурусы и взаимодействие с другими словарями. Часть 1: Тезаурус для информационного поиска») была опубликована в августе 2011 г. С принятием этого стандарта были отменены ранее действовавшие стандарты ISO 2788 и ISO 5964, положения которых были пересмотрены и обновлены. В 2015 г. первая часть ISO 25964 была введена в российскую систему стандартизации как *ГОСТ Р 7.0.91–2015 «СИБИД. Тезаурусы для информационного поиска»* [15].

Новые стандарты устанавливают рекомендации для развития и ведения не только ИПТ, их можно применить и к другим словарям, используемым в процессе поиска информации по всем видам информационных ресурсов. Международный стандарт может быть применен как к однопользовательным, так и к многопользовательным тезаурусам.

В российскую версию не вошли разделы, касающиеся построения многоязычных тезаурусов, поскольку в России эти вопросы сегодня регулируются стандартом ГОСТ 7.24–2007. В стандартах представлены определения основных терминов и концепций, используемых при создании ИПС, определены типы связей между терминами и концепциями ИПС, также даны рекомендации по созданию и ведению, оформлению и представлению тезауруса.

Структура тезауруса определена как модель данных, изначально разработанная в британском стандарте и представленная как диаграмма классов унифицированного языка моделирования UML (*Unified Modeling Language*), где каждое понятие формируется совокупностью атрибутов экземпляров классов. Приведенная в стандартах модель структуры данных может служить основой для создания тезауруса. Модель данных содержит пять основных классов: тезаурус, массив тезауруса, концепцию тезауруса, термин тезауруса и примечание. Атрибуты для каждого класса и ассоциации классов отражают все функции тезаурусов, определенные в тексте стандартов. Модель сопровождается четкими пояснительными примечаниями. Представление модели в форме UML строго определяет структуру и функции тезауруса, что позволяет ее использовать для создания спецификации по созданию и использованию программного обеспечения ИПС, исключая неопределенности.

С учетом того обстоятельства, что некоторые пользователи могут быть не знакомы со спецификацией UML, в стандартах также приведено табличное представление модели данных ИПТ. Элементы данных и атрибуты диаграммы UML аналогично представлены в таблицах и идентичны по составу, хотя их отношения в табличном представлении сложнее проследить. Кроме того, таблицы не отражают некоторые детали, связанные с нотацией UML.

ISO 25964-2 и межсистемная совместимость тезаурусов и других KOS

Вторая часть международного стандарта *ISO 25964-2 «Information and documentation – Thesauri and interoperability with other vocabularies – Part 2: Interoperability with other vocabularies»* (ISO 25964-2 «Информация и документация. Тезаурусы и взаимодействие с другими словарями. Часть 2: Взаимодействие с другими словарями») была разработана и введена в действие в 2013 г. В стандарте описаны методы сопоставления структуры и элементов тезаурусов и других ти-

⁷ Model. The official development website for BS 8723: Structured Vocabularies for Information Retrieval. Available from: <https://schemas.bs8723.org/model/> (accessed: 25.04.2025).

пов словарей, которые обычно используются для поиска информации. Основное внимание уделяется принципам моделирования и практике сопоставления понятий, представленных в контролируемых словарях различных видов. Стандарт дает рекомендации по созданию и поддержанию соответствий между несколькими тезаурусами или между тезаурусами и другими типами словарей. Область применения включает взаимодействие со схемами классификаций, таксономиями, словарями предметных рубрик, онтологиями, терминосистемами, авторитетными файлами и словарями синонимических рядов. Стандарт регламентирует методы интеграции структурных компонентов тезаурусов и иных лексических систем, используемых в информационном поиске. Представлены принципы моделирования связей между понятиями, зафиксированными в регламентированных словарях, и их практическое согласование. Стандарт содержит руководства по установлению и поддержанию семантических соответствий как между различными тезаурусами, так и между тезаурусами и другими типами терминологических систем. Сфера применения стандарта охватывает взаимодействие:

- с классификационными схемами и таксономиями;
- системами предметных рубрик;
- онтологиями и терминологическими базами;
- нормативными базами данных (авторитетные файлы);
- словарями синонимов и близкородственных понятий.

Стандарт содержит также рекомендации по разработке сопоставительных таблиц, которые указывают соответствия элементов различных схем и словарей.

Разработка стандарта велась в тесном сотрудничестве экспертов рабочих групп W3C SKOS и ISO 25964, в результате чего модели данных, представленные в этих стандартах, в значительной степени совместимы. Некоторые различия связаны с различиями в области применения. SKOS ориентирован исключительно на веб-приложения, использующие все виды систем организаций знаний (KOS), перечень которых включает схемы классификации, не включенные в ISO 25964. В то же время для ISO 25964 широко определена область применения (включая веб-приложения), но только относительно правил формирования тезаурусов. Несмотря на различия в областях применения, при разработке стандартов была достигнута хорошая согласованность между компонентами моделей данных; кроме того, для SKOS существует

механизм гармонизации различий, возникающих в процессе развития различных KOS, включая модель ISO 25964. Элементы, отсутствующие в базовой модели SKOS (например, сложная эквивалентность, англ. *compound equivalence*, представленная в ISO 25964), могут быть включены в расширенную модель SKOS-XL, которая может дополняться по необходимости.

Развитие российских ГОСТов на основе ISO 25964

В российскую систему стандартизации ISO 25964 вошел как ГОСТ Р 7.0.106–2024 «СИБИД. Взаимодействие тезаурусов и других словарей», модифицированный по отношению к международному.

В процессе разработки двух частей стандарта ISO 25964 в период 2011–2013 гг. были актуализированы принципы формирования ИПТ, сегодня эти стандарты представляют современную нормативную базу тезаурусов. Вместе с тем действовавшие на тот период стандарты по методам поискового индексирования ISO 999:1996 и ISO 5963:1985 были разработаны в то время, когда Интернет только начинал свое развитие как глобальная информационно-коммуникационная среда, и их положения относятся преимущественно к ручным методам индексирования печатных изданий.

Перспективы нормативной трансформации: проект новой редакции ISO 999

В 2023 г. в ИСО/ТК 46 «Информация и документация» было принято решение об организации рабочей группы РГ 12 «Индексы и индексирование» (ISO/TC 46/SC 9/WG 12 “Indexes and indexing”) для разработки проекта новой, расширенной редакции стандарта ISO 999, объединяющего положения ISO 999:1996 и ISO 5963:1985, с последующей их отменой. В рабочую группу вошли представители семи стран, включая Россию. От России в качестве эксперта в работе группы участвует автор статьи. Руководителем группы был избран представитель США доктор (Ph. D.) Николас Милиарас (Nicholas Miliaras), который являлся руководителем разработки действующего стандарта ISO 999:1996.

Проект обновленного стандарта определяет базовые принципы индексирования, разработанные с учетом новых требований, связанных с появлением новых видов информационных ресурсов и развитием коммуникационных технологий в сети Интернет. Положения, установленные в стандарте, относятся как к словарям, определяемым системой стандартов ISO 25964, так и к си-

стемам индексирования SKOS. Они относятся к формированию всех типов индексов независимо от того, создаются ли те вручную или с помощью компьютерных методов, в том числе средствами искусственного интеллекта. Представленные в стандарте рекомендации могут быть применены к любым из существующих методов индексирования, таким как, например, координатное индексирование библиотечных ресурсов, выбор ключевых слов при подготовке публикаций, SEO-оптимизация сайтов или индексирование цитирования. Кроме того, в стандарте даны рекомендации по организации самого процесса индексирования и контролю качества формируемых индексов.

Проект стандарта определяет применение различных видов индексов, таких как тематические и авторские указатели, указатели имен, заголовков, номерные, кодированные, а также фактографические индексы. Применяемые индексы могут включать записи по различным категориям, включая имена (личные, корпоративные, географические), технические термины, темы, названия произведений, первые строки стихотворений, цитаты, сокращения, цифры, даты и т.д. Кроме того, применяемые индексы могут быть:

- однократными (закрытый индекс) или периодически обновляемыми (открытый индекс);
- доступными для визуального восприятия человеком (отображаемый индекс) или доступными только для поиска компьютерными средствами (неотображаемый индекс);
- определяемыми путем интеллектуального анализа человеком или техническими средствами с помощью методов искусственного интеллекта. Также возможно создание индексов комбинированными методами человеческого и компьютерного анализа.

Сформированные индексы могут объединяться в единую последовательность или разделяться на отдельные группы, а также могут составлять иерархические структуры.

При разработке новой редакции стандарта был учтен опыт, накопленный в других международных и национальных организациях по стандартизации. В качестве примера можно назвать документы *Guidelines for Multilingual Thesauri* («Руководство для многоязычных тезаурусов»)⁸ и *Guidelines for Subject Access in National Bibliographies* («Руководство по обеспечению тематического доступа в национальной библиографии») [16], разработанные

в IFLA в секции по классификации и индексированию. В связи с этим также следует отметить работу *American Society for Indexing, ASI* (Американское общество индексирования), опубликовавшего в 2015 г. свои рекомендации по индексированию⁹, касающиеся новых видов электронных публикаций и дополняющие такой важный для издателей США и других англоязычных стран нормативный документ, как *The Chicago Manual of Style* («Чикагское руководство по стилю»)¹⁰.

МЕТОДОЛОГИЧЕСКИЙ РАЗРЫВ И ВЫЗОВЫ СОВМЕСТИМОСТИ

Рассмотренные примеры показывают, что современная нормативная база поискового индексирования в России вступает в противоречие с ключевыми требованиями цифровой среды. Несмотря на частичную адаптацию международных стандартов (например, внедрение ГОСТ Р 7.0.91–2015 на основе ISO 25964-1), системной интеграции таких моделей, как RDF и SKOS, до сих пор не произошло. Как следствие, отечественные KOS остаются неинтероперабельными, непрозрачными для автоматических систем обработки информации и изолированными от глобального информационного обмена.

Отставание проявляется не только в технических аспектах, но и на методологическом уровне. Действующие стандарты продолжают опираться на терминоцентричную парадигму, разработанную в эпоху бумажной документации и ручного индексирования. В них отсутствует понятийная модель, необходимая для автоматической обработки концептов и построения семантически согласованных словарей. Это противоречит самой логике цифровой трансформации, основанной на онтологически управляемых системах и распределенных средах.

Попытки частичной модернизации, такие как введение отдельных ГОСТов, не компенсируют отсутствия единой концептуальной рамки, обеспечивающей преемственность между классификациями, онтологиями, авторитетными файлами и тезаурусами. Недостаточное внимание к межсловарной совместимости ограничивает возможности транскрипционного поиска, снижает цитируемость ресурсов и затрудняет участие в международных инфраструктурных инициативах.

⁹ Best Practices for Indexing. 1st ed. Tempe, AZ: American Society for Indexing; 2015. 84 p. Available from: <https://asindexing.org/wp-content/uploads/2014/04/ASI-BestPracticesIndexing2015-linked-updated-copyright-v3-02-26-16.pdf> (accessed: 25.04.2025).

¹⁰ The Chicago Manual of Style. 17th ed. Chicago, IL: The University of Chicago Press; 2017. 1176 p. Available from: <https://archive.org/details/chicagomanualofstyle0017unse> (accessed: 02.07.2024).

⁸ Guidelines for Multilingual Thesauri. The Hague: IFLA Headquarters; 2009. 30 p. (IFLA Professional Reports. No. 115). ISBN 978-90-77897-35-5. ISSN 0168-1931. Available from: <https://ifap.ru/library/book411.pdf> (accessed: 25.04.2025).

Таким образом, дальнейшее сохранение фрагментарного подхода к нормативному регулированию не только тормозит развитие отечественной информационной инфраструктуры, но и угрожает ее маргинализацией в условиях растущей зависимости от глобальных цифровых экосистем. Требуется переход от изолированных обновлений к системной модернизации нормативной базы, ориентированной на принципы совместимости, открытости и концептуальной прозрачности.

РЕКОМЕНДАЦИИ ПО ОБНОВЛЕНИЮ НОРМАТИВНОЙ БАЗЫ И ПРОФЕССИОНАЛЬНОЙ ПРАКТИКИ

Обновление нормативной базы поискового индексирования должно учитывать современные требования цифровой среды и обеспечить переход от устаревших принципов ручного индексирования к онтологическим моделям, совместимым с международными стандартами ISO 25964 и SKOS. На основе проведенного анализа целесообразно выделить несколько направлений, по которым следует разработать и реализовать соответствующие меры.

Пересмотр и отмена устаревших нормативных документов

Действующие в настоящее время стандарты ГОСТ 7.24–80 и ГОСТ 7.25–80 основываются на положениях международных стандартов ISO 5964 и ISO 2788, которые были отменены в 2011 г. в связи с принятием нового стандарта ISO 25964. С учетом их устаревания необходимо инициировать официальную отмену указанных документов.

Отмена ГОСТ 7.24 повлечет за собой необходимость пересмотра ГОСТ Р 7.0.91–2015, в котором отсутствуют положения, касающиеся построения многоязычных тезаурусов. Эти положения изложены в части 2 международного стандарта ISO 25964:2013 и должны быть включены в российскую версию при следующем пересмотре.

Гармонизация с международными моделями индексирования

Современные требования цифровой среды, включая активное распространение информационных ресурсов в Интернете и использование технологий семантической обработки, обуславливают необходимость пересмотра принципов индексирования и формализации семантических связей между понятиями. В связи с этим необходимо обеспечить совместимость национальных стандартов с моделью представления данных RDF и стандартом SKOS (W3C, 2009), который уже ак-

тивно применяется в международной библиотечной и издательской практике.

Стандарт SKOS представляет собой концептуальную модель, в которой основными элементами являются не термины, а абстрактные понятия, описываемые через термины и их свойства. Он обеспечивает формализацию и интероперабельность структур знаний, включая иерархические и ассоциативные связи, аналогичные используемым в ISO 2788 и ISO 5964, но адаптированные для цифровой среды. Совместимость с SKOS позволит интегрировать отечественные тезаурусы в глобальные цифровые платформы и обеспечить автоматизированную обработку понятийных связей.

Разработка национального руководства по интероперабельности KOS

С учетом растущего числа различных словарей, используемых в информационном поиске (тезаурусы, классификационные схемы, онтологии, авторитетные файлы), необходим нормативный документ, регламентирующий правила их сопоставления и совместного использования. В этом контексте целесообразно адаптировать положения ISO 25964-2:2013 в части моделирования связей между понятиями, зафиксированными в различных типах словарей.

Новый национальный документ должен включать рекомендации по установлению соответствий между терминами разных тезаурусов, а также между тезаурусами и другими системами организации знаний. При этом следует учитывать как структуру самих KOS, так и механизмы их технической реализации (включая SKOS, OWL и RDF).

Участие в международной стандартизации

Развитие нормативной базы должно сопровождаться более активным участием российских специалистов в деятельности международных организаций по стандартизации. Показательным примером может служить участие автора настоящей статьи в составе рабочей группы ISO/TC 46/SC 9/WG 12, разрабатывающей проект новой версии стандарта ISO 999, объединяющего положения ISO 999:1996 и ISO 5963:1985. Это позволяет учитывать специфику российской практики и способствовать ее согласованию с глобальной нормативной архитектурой.

Подготовка к внедрению новой редакции ISO 999

В 2023 г. в рамках ISO/TC 46 была инициирована разработка проекта новой, расширенной редакции международного стандарта ISO 999, которая объединяет положения действующих ISO 999:1996 и ISO 5963:1985 и предназначе-

на для индексирования всех видов информационных ресурсов в цифровой среде с учетом современных технологий, включая методы искусственного интеллекта. Введение этого стандарта в российскую систему СИБИД планируется в 2027 г. в идентичном или модифицированном виде. В связи с этим требуется подготовка методических и организационных условий для его внедрения, включая пересмотр национальных стандартов ГОСТ 7.59–2003 и ГОСТ Р 7.0.66–2010.

Повышение квалификации специалистов и трансформация практики индексирования

Обновление нормативной базы должно сопровождаться мерами по повышению квалификации специалистов, работающих в библиотечной, информационной и издательской сферах. В условиях цифровой трансформации индексирования необходимо обеспечить подготовку профессионалов, владеющих современными инструментами семантической обработки, онтологического моделирования и работы со структурами типа SKOS. Также следует разработать учебные и методические материалы, обеспечивающие единое понимание новых принципов и механизмов индексирования.

ЗАКЛЮЧЕНИЕ

Современная информационная среда предъявляет радикально иные требования к организации знаний, чем те, на которых основывались отечественные стандарты в области поискового индексирования в конце XX в. Сохраняющаяся в России нормативная база, несмотря на формальное обновление отдельных документов, продолжает функционировать на принципах, унаследованных от устаревших международных стандартов ISO 2788 и ISO 5964. В условиях цифровизации, перехода к автоматизированным

системам обработки данных и роста требований к семантической совместимости такая ситуация приводит не только к снижению эффективности информационного поиска, но и к институциональной изоляции отечественных информационных ресурсов. Проведенный анализ показал, что основной проблемой является не отсутствие технических решений, а нормативно-концептуальное несоответствие действующих стандартов современным моделям организации знаний, прежде всего основанным на RDF и SKOS. Преодоление этого разрыва требует не фрагментарных мер, а системной трансформации, охватывающей законодательную, технологическую и образовательную сферы.

Предложенные в статье рекомендации отражают стратегическую необходимость отказа от архаичных представлений об индексировании и перехода к межоперабельным, машиночитаемым и концептуально устойчивым моделям. Внедрение положений ISO 25964 и SKOS, пересмотр ГОСТов, подготовка к принятию новой редакции ISO 999, развитие профессиональных компетенций – все это неотъемлемые элементы политики модернизации. Следует подчеркнуть, что данная модернизация не является уступкой внешним требованиям, а обусловлена внутренней логикой развития самой информационной отрасли. Без интеграции в международную архитектуру стандартов и знаний российские информационные системы рискуют оказаться на периферии глобального информационного пространства, утратив как функциональность, так и научную видимость. Таким образом, обновление нормативной базы индексирования должно рассматриваться не как техническая процедура, а как необходимое условие сохранения и развития национального информационного суверенитета в условиях цифровой трансформации.

КОНФЛИКТ ИНТЕРЕСОВ

Автор заявляет об отсутствии конфликта интересов.

CONFLICT OF INTERESTS

The author declares no relevant conflict of interests.

СПИСОК ЛИТЕРАТУРЫ

1. Dextre Clarke S.G., Lei Zeng M. From ISO 2788 to ISO 25964: The evolution of thesaurus standards towards interoperability and data modeling. *Information Standards Quarterly*. 2012;24(1):20–26. <https://doi.org/10.3789/isqv24n1.2012.04>
2. Тимошенко И.В. Современные тенденции развития методов и нормативной базы индексирования библиотечных информационных ресурсов. *Научные и технические библиотеки*. 2024;(10):102–122. <https://doi.org/10.33186/1027-3689-2024-10-102-122>

3. Белоозеров В.Н., Антошкова О.А. Разработка национального стандарта на сопоставление информационно-поисковых языков. В кн.: *Научное издание международного уровня – 2019: стратегия и тактика управления и развития: материалы 8-й Междунар. науч.-практ. конф. (Москва, 23–26 апреля 2019 г.)*. Екатеринбург: Изд-во Урал. ун-та; 2019. С. 35–43. <https://doi.org/10.24069/konf-23-26-04-2019.04>
4. Солтон Дж. *Динамические библиотечно-информационные системы*. Пер. с англ. М.: Мир; 1979. 557 с.
5. Довбенко М.А., Уманский А.М. *Информационно-поисковые языки автоматизированных систем по стандартизации*. М.: Изд-во стандартов; 1981. 239 с.
6. Гендина Н.И. *Лингвистическое обеспечение автоматизированных библиотечных систем*. Алма-Ата: Гылым; 1991. 204 с.
7. Черный А.И. *Введение в теорию информационного поиска*. М.: Наука; 1975. 238 с.
8. Лукашевич Н.В. *Тезаурус в задачах информационного поиска*. М.: Изд-во МГУ; 2011. 512 с.
9. Антопольский А.Б., Вайсберг А.М., Зарувинская Л.А. Принципы создания и функционирования автоматизированной системы ведения информационных языков. *Научно-техническая информация. Серия 2. Информационные процессы и системы*. 1976;(6):8–12.
10. Тимошенко И.В. *Стандартизация информационных видов деятельности. Система стандартов по информации, библиотечному и издательскому делу*. М.: Знание-М; 2023. 232 с.
11. Gibbs J. Standards for indexing: Evolution since 1989. *Information Standards Quarterly*. 2015;27(2-3):41–45. Available from: <https://www.niso.org/sites/default/files/stories/2017-08/Standards%20for%20Indexing.pdf> (accessed: 24.07.2025).
12. Riva P., Le Boeuf P., Žumer M. *IFLA Library Reference Model: A Conceptual Model for Bibliographic Information*. Den Haag: IFLA; 2017. 101 p. Available from: <https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017.pdf> (accessed: 25.04.2025).
13. Matthews B., Miles A. *Review of RDF Thesaurus Work: A Review and Discussion of RDF Schemas for Thesauri*. SWAD-Europe Thesaurus Activity. Deliverable 8.2. Available from: <http://www.w3.org/2001/sw/Europe/reports/thes/8.2/draft01.html> (accessed: 25.04.2025).
14. Miles A., Bechhofer S., editors. *SKOS Simple Knowledge Organization System Reference: W3C Recommendation 18 August 2009*. W3C. Available from: <http://www.w3.org/TR/2009/REC-skos-reference-20090818/> (accessed: 25.04.2025).
15. Тимошенко И.В. Нормативная база информационно-поисковых тезаурусов: история и современные тенденции развития. *Вестник РГТУ. Серия: Информатика. Информационная безопасность. Математика*. 2024;(2):8–25. <https://doi.org/10.28995/2686-679X-2024-2-8-25>
16. Jahns Y., editor. *Guidelines for Subject Access in National Bibliographies*. IFLA Working Group on Guidelines for Subject Access by National Bibliographic Agencies. Berlin: De Gruyter Saur; 2012. 109 p. (IFLA Series on Bibliographic Control. Vol. 45).

REFERENCES

1. Dextre Clarke S.G., Lei Zeng M. From ISO 2788 to ISO 25964: The evolution of thesaurus standards towards interoperability and data modeling. *Information Standards Quarterly*. 2012;24(1):20–26. <https://doi.org/10.3789/isqv24n1.2012.04>
2. Timoshenko I.V. The current trends in the development of methods and regulatory framework for indexing library information resources. *Nauchnye i tekhnicheskie biblioteki = Scientific and Technical Libraries*. 2024;(10):102–122. (In Russ.). <https://doi.org/10.33186/1027-3689-2024-10-102-122>
3. Beloozerov V.N., Antoshkova O.A. Development of a national standard for the comparison of information retrieval languages. In: *World-Class Scientific Publication – 2019: Strategy and Tactics of Management and Development. Proc. 8th Int. Sci.-Pract. Conf. (Moscow, April 23–26, 2019)*. Ekaterinburg: Ural University Publ.; 2019, pp. 35–43. (In Russ.). <https://doi.org/10.24069/konf-23-26-04-2019.04>
4. Salton G. *Dynamic Information and Library Processing*. Englewood Cliffs, NJ: Prentice-Hall; 1975. 523 p. (Russ. ed.: Salton G. *Dinamicheskie bibliotechno-informatsionnye sistemy*. Moscow: Mir Publ., 1979. 557 p.).
5. Dovbenko M.A., Umaniskii A.M. *Information Retrieval Languages of Automated Systems for Standardization*. Moscow: Izdatelstvo standartov; 1981. 239 p. (In Russ.).
6. Gendina N.I. *Linguistic Support of Automated Library Systems*. Alma-Ata: Gylm; 1991. 204 c. (In Russ.).
7. Chernyi A.I. *Introduction to Information Retrieval Theory*. Moscow: Nauka; 1975. 238 p. (In Russ.).
8. Lukashevich N.V. *Thesaurus in Information Retrieval Tasks*. Moscow: Lomonosov Moscow State University Publ.; 2011. 512 p. (In Russ.).
9. Antopol'skii A.B., Vaisberg A.M., Zaruvin'skaya L.A. Principles of creation and operation of an automated system for maintaining information languages. *Nauchno-tekhnicheskaya informatsiya. Seriya 2. Informatsionnye protsessy i sistemy*. 1976;(6):8–12. (In Russ.).

10. Timoshenko I.V. *Standardization of Information Activities. System of Standards for Information, Library and Publishing*. Moscow: Znanie-M; 2023. 232 p. (In Russ.).
11. Gibbs J. Standards for indexing: Evolution since 1989. *Information Standards Quarterly*. 2015;27(2-3):41–45. Available from: <https://www.niso.org/sites/default/files/stories/2017-08/Standards%20for%20Indexing.pdf> (accessed: 24.07.2025).
12. Riva P., Le Boeuf P., Žumer M. *IFLA Library Reference Model: A Conceptual Model for Bibliographic Information*. Den Haag: IFLA; 2017. 101 p. Available from: <https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017.pdf> (accessed: 25.04.2025).
13. Matthews B., Miles A. *Review of RDF Thesaurus Work: A Review and Discussion of RDF Schemas for Thesauri*. SWAD-Europe Thesaurus Activity. Deliverable 8.2. Available from: <http://www.w3.org/2001/sw/Europe/reports/thes/8.2/draft01.html> (accessed: 25.04.2025).
14. Miles A., Bechhofer S., editors. *SKOS Simple Knowledge Organization System Reference: W3C Recommendation 18 August 2009*. W3C. Available from: <http://www.w3.org/TR/2009/REC-skos-reference-20090818/> (accessed: 25.04.2025).
15. Timoshenko I.V. The normative base of information retrieval thesauri. History and current development trends. *Vestnik RGGU. Seriya: Informatika. Informatsionnaya bezopasnost'. Matematika = RSUH/RGGU Bulletin. Series: Information Science. Information Security. Mathematics*. 2024;(2):8–25. (In Russ.). <https://doi.org/10.28995/2686-679X-2024-2-8-25>
16. Jahns Y., editor. *Guidelines for Subject Access in National Bibliographies*. IFLA Working Group on Guidelines for Subject Access by National Bibliographic Agencies. Berlin: De Gruyter Saur; 2012. 109 p. (IFLA Series on Bibliographic Control. Vol. 45).

ИНФОРМАЦИЯ ОБ АВТОРЕ

Игорь Владимирович Тимошенко, кандидат технических наук, доцент ВАК, доцент, Московский государственный институт культуры, ответственный секретарь ТК 191 «Научно-техническая информация, библиотечное и издательское дело, управление документами», г. Москва, Российская Федерация; <https://orcid.org/0000-0003-2995-1662>; e-mail: timigor@gpntb.ru

INFORMATION ABOUT THE AUTHOR

Igor V. Timoshenko, Cand. Sci. (Eng.), Associated Professor HAC, Associated Professor, Moscow State Institute of Culture, Executive Secretary of GOST RU/TC 191 “Scientific and Technical Information, Librarian and Publishing Business, Documents Management”, Moscow, Russian Federation; <https://orcid.org/0000-0003-2995-1662>; e-mail: timigor@gpntb.ru

Поступила в редакцию / Received 28.04.2025

Поступила после рецензирования / Revised 29.06.2025

Принята к публикации / Accepted 24.07.2025