

БИБЛИОТЕЧНЫЕ КАТАЛОГИ И ИНФОРМАЦИОННО-ПОИСКОВЫЕ СИСТЕМЫ

УДК 025.4.036 + 004:025

<https://doi.org/10.33186/1027-3689-2024-11-103-120>

Формирование связанных данных в модели Единого открытого архива информации ГПНТБ России (ЕОАИ) с использованием многоязычного тезауруса GEMET

М. В. Гончаров^{1,2}, К. А. Колосов^{1, 2}

¹ГПНТБ России, Москва, Российская Федерация

²Московский государственный лингвистический университет,
Москва, Российская Федерация

^{1, 2}goncharov@gpntb.ru

^{1, 2}kolosov@gpntb.ru

Аннотация. Практическое применение связанных данных в библиотечных проектах означает, с одной стороны, представление их в формате RDF для последующего использования и обработки информационными системами, а с другой – обогащение информации, выдаваемой пользователям поисковых порталов. В данном исследовании мы рассматривали вопросы, относящиеся ко второму направлению, то есть обогащению информации, формируемой при выводе результатов поиска. В качестве базового набора данных для построения связей нами использовался многоязычный тезаурус GEMET. Портал Единого открытого архива информации ГПНТБ России (ЕОАИ) разрабатывается с учётом современных технологий и тенденций в области использования связанных данных, обзор которых приводится в начале статьи. Отмечается, что, несмотря на обилие наборов данных, доступных на ресурсе облака связанных открытых данных (Linked Open Data Cloud), порталы библиотечных проектов, применяющие технологию связанных данных, используют лишь отдельные пересекающиеся между собой наборы данных.

Авторами статьи был проведён анализ записей в электронном каталоге ЕОАИ и электронном каталоге базы данных «Экология: наука и технология» с целью автоматизированного формирования связей между индексами УДК и предметными рубриками библиографических записей с концептами тезауруса GEMET. Полученные результаты показывают, что в обеих рассмотренных базах данных около 20% записей содержат предметные рубрики и индексы УДК, которые могут быть связаны с концептами тезауруса GEMET.

Наибольший интерес для практического использования на пользовательском портале представляют связи концептов тезауруса GEMET со страницами DBpedia в сети Интернет. Авторы предлагают дополнить эти связи ссылками на страницы российской онлайн-энциклопедии РУВИКИ. Эти дополнительные связи были сделаны в рабочей базе данных для верхних уровней УДК и в дальнейшем будут прописаны для всех используемых концептов тезауруса GEMET.

Статья подготовлена в рамках Государственного задания ГПНТБ России № 075-01235-23-01 по теме «Информационное обеспечение научных исследований учёных и специалистов на базе Открытого архива ГПНТБ России как системы агрегации научных знаний» (FNEG-2022-0003) на 2022–2024 гг. № 1021062311369-1-1.2.1;5.8.2;5.8.3.

Ключевые слова: связанные данные, LOD, открытые архивы, Единый открытый архив информации, RDF, BIBFRAME, SKOS, GEMET

Для цитирования: Гончаров М. В., Колосов К. А. Формирование связанных данных в модели Единого открытого архива информации ГПНТБ России (ЕОАИ) с использованием многоязычного тезауруса GEMET // Научные и технические библиотеки. 2024. № 11. С. 103–120. <https://doi.org/10.33186/1027-3689-2024-11-103-120>

LIBRARY CATALOGS AND INFORMATION RETRIEVAL SYSTEMS

UDC 025.4.036 + 004:025

<https://doi.org/10.33186/1027-3689-2024-11-103-120>

Generating linked data in the model of RNPLS&T Single Open Information Archive (SOIA) using GEMET multilingual thesaurus

Mikhail V. Goncharov^{1, 2} and Kirill A. Kolosov^{1, 2}

¹*Russian National Public Library for Science and Technology,
Moscow, Russian Federation*

²*Moscow State Linguistic University, Moscow, Russian Federation*

^{1, 2}goncharov@gpntb.ru

^{1, 2}kolosov@gpntb.ru

Abstract. The practical use of linked data in library projects implies their representation in the RDF format for further use and processing in information systems, and, on the other hand, enrichment of information delivered to the search portal users. In this study, the authors examine the aspects of information enrichment for search outcome. They used GEMET multilingual thesaurus as the basic data set for linking. The RNPLS&T Single Open Information Archive (SOIA) portal has been built based on advanced technologies and trends in using linked data. These trends are also reviewed. Despite plenty of data sets available via the Linked Open Data Cloud, the library projects embracing the linked data technology use just few overlapping data sets.

To build the links between UDC codes and subject headings with GEMET concepts, the authors analyzed the entries in the e-catalogs of SOIA and “Ecology: Science and Technology” database. The analysis results demonstrate that approximately 20% entries comprise subject headings and UDC codes that might be linked to GEMET thesaurus concepts.

The links between GEMET concepts and Internet-based DBpedia are seen as the most promising option. The authors suggest to supplement these links with the references to Ruwiki (Russian online encyclopedia). These additional links were established for UDC upper levels in the operational database and will be further created for every GEMET active concept.

The article is prepared within the framework of the Government Order No. 075-01235-23-01, theme No. 1021062311369-1-1.2.1;5.8.2;5.8.3 “Information support of research by scientists and specialists on the basis of RNPLST Open Archive as the system of scientific knowledge aggregation” (FNEG-2022-0003) for 2022–2024.

Keywords: linked data, LOD, open archive, single open information archive, RDF, BIBFRAME, SKOS, GEMET

Cite: Goncharov M. V., Kolosov K. A. Generating linked data in the model of RNPLS&T Single Open Information Archive (SOIA) using GEMET multilingual thesaurus // Scientific and technical libraries. 2024. No. 11, pp. 103–120. <https://doi.org/10.33186/1027-3689-2024-11-103-120>

Введение

Технологии связанных открытых данных (LOD) стали всё активнее внедряться в библиотеках, архивах и музеях, расширяя традиционные модели информационных систем этих организаций. Документы и метаданные, которыми обладают эти организации, особенно библиотеки, имеют высокий уровень структурирования. Ещё в 2004 г. Консорциум Всемирной паутины (W3C) рекомендовал библиотекам публиковать свои данные с использованием технологий семантической паутины для улучшения их цифрового воздействия и повышения полезности [1]. Среди преимуществ использования этой модели можно выделить следующие:

- улучшенную видимость коллекций;
- повторное использование данных;
- упрощение процессов каталогизации;
- сотрудничество с другими сообществами, генерирующими данные;
- снижение затрат для учреждений.

Данные, опубликованные как связанные данные в формате RDF, предназначены для машинной обработки и использования приложениями, которые с их помощью создают новые ресурсы и услуги посредством массовых загрузок или выборочных запросов данных с использованием точек доступа SPARQL или других технологий [2].

Как отмечается в [3], OCLC и крупные национальные библиотеки возглавляют важные проекты в области связанных данных, в то время как

большинство небольших библиотек лишь наблюдают за прототипами, которые демонстрируют преимущества публикации данных в виде связанных данных, а многие библиотекари не могут представить, как станет выглядеть будущий семантический каталог и как он будет работать.

Наиболее известным примером использования связанных данных в библиотеках является библиографическая структура BIBFRAME – модель, предложенная как эволюция формата MARC 21 в среду Семантического веба и связанных открытых данных (LOD). Словарь, используемый в BIBFRAME, основан на классах и свойствах RDF для обмена данными в интернете. Эта структура расширяет структуру ссылок интернета за счёт использования триплетов или предложений RDF и деклараций о веб-ресурсах в форме выражений «субъект – предикат – объект», каждый из которых идентифицируется с помощью уникального идентификатора ресурсов (URI). Однако, как отмечается [3], очень нелегко преобразовать огромный объём данных, каталогизированных за сорок лет, в связанные данные. Библиотечные онтологии и словари сложны, и для завершения преобразования требуется значительное количество времени. По мнению авторов статьи [2], BIBFRAME не будет единственной библиографической структурой данных. Она будет существовать с другими структурами данных, такими как Schema.org, Europeana Data Model (EDM), и, вероятно, некоторыми другими. Отвечая на обеспокоенность относительно различных концептуальных моделей связанных данных в библиотечной сфере, авторы публикации [4] поясняют, что они сравнивают и ищут конвергенции, которые гарантируют совместимость между четырьмя моделями: FRBR, FRBRoo (объектно-ориентированная версия FRBR), BIBFRAME и EDM.

Как отмечают отдельные авторы, изменения в отношении интерфейсов поиска и функциональных возможностей библиографических информационных систем, включая библиотечные каталоги, с внедрением технологии связанных данных не были столь значительными, как ожидалось [3].

Преимущества использования модели связанных данных для пользователей библиотечных порталов

Для конечного пользователя преимуществом использования модели связанных открытых данных является обогащение информации, получаемой в результате обработки поисковых запросов. Публикация наборов данных как связанных открытых данных (LOD) открывает мно-

жество возможностей исследовать и взаимодействовать с каталогом в новых контекстах, улучшая возможности обнаружения и связи с другими коллекциями. Семантическое обогащение библиографических записей знаниями LOD отвечает более сложным информационным потребностям пользователей для расширенного поиска в интуитивно понятной среде.

Ресурс Linked Open Data Cloud [5] содержит наборы связанных открытых данных, включая информацию о точках доступа для запросов на языке SPARQL, а также позволяет выгружать наборы данных для последующего их использования. По состоянию на конец августа 2024 г., данный ресурс содержит 1346 наборов связанных данных. Данные могут быть использованы в соответствии с положениями лицензии Creative Commons Attribution 4.0 (<https://creativecommons.org/licenses/by/4.0>). Несмотря на обилие наборов данных, доступных на указанном ресурсе, порталы библиотечных проектов, применяющие технологию связанных данных, используют лишь отдельные пересекающиеся между собой наборы данных. В табл. 1 представлены наборы связанных данных, используемые в следующих библиотечных проектах:

Virtual Polygraph Library Ignacio Larramendi (Biblioteca Virtual de Polígrafos) – **BVP**;

Open Data Catalogue of the Miguel de Cervantes Virtual Library (Catálogo de Datos Abiertos de la Biblioteca Virtual Miguel de Cervantes) – **DBVMC**;

Bibliographic Data Portal of the National Library of Spain (Portal de Datos Bibliográficos de la Biblioteca Nacional de España) – **DBNE**;

DataBnF of the National Library of France – **DBNF**.

Таблица 1

**Наборы данных, используемые в проектах связанных данных
некоторых крупных библиотек**

Проект	VIAF	ISNI	GeoNames	Wikidata	DBpedia	LCSH
BVP	+		+	+	+	+
DBVMC	+	+		+		
DBNE	+	+		+	+	+
DBNF	+	+	+	+	+	

Отметим особенности наборов данных, приведённых в табл. 1.

VIAF (Virtual International Authority File) – это совместный проект 20 национальных библиотек, реализованный и размещённый OCLC,

который фактически объединяет несколько авторитетных файлов имён авторов, собранных из библиотек, архивов, музеев, в единый авторитетный сервис имён. Участником проекта от России является Российская национальная библиотека (РНБ). Информация по каждому автору содержит формы написания на разных языках, связь с индексом ISNI, ссылку на страницу в Википедии (если имеется), сведения о некоторых произведениях автора, годы жизни (опционально).

ISNI (International Standard Name Identifier – ISO 27729) – Международный идентификатор стандартных наименований – метод для уникальной идентификации создателей и правообладателей различных видов медиаконтента, таких как книги, телевизионные передачи, газетные статьи, концерты, фильмы. Регистрация имени ISNI позволяет присвоить уникальный идентификатор каждому создателю творческого содержания (например, авторскому псевдониму) в виде уникального номера Творческого паспорта. ISNI используется ведущими поисковыми системами, библиотеками и архивами при обмене информационными каталогами, для более точного поиска информации в сети Интернет и в базах данных, а также может помочь в управлении правами интеллектуальной собственности на международном уровне и в цифровой среде.

GeoNames – географическая база данных, доступная через различные веб-сервисы по условиям лицензии Creative Commons. Одна из крупнейших и наиболее широко используемых баз геоданных в части названий и координат различных мест.

Wikidata – бесплатная и открытая база знаний, которую могут читать и редактировать как люди, так и машины. Wikidata выступает в качестве центрального хранилища структурированных данных родственных проектов Wikimedia, включая Wikipedia, Wikivoyage, Wiktionary, Wikisource и др.

DBpedia – это совместная работа по извлечению структурированной информации из Википедии и предоставлению этой информации в интернете. Позволяет формировать сложные запросы к Википедии и связывать другие наборы данных в интернете с данными Википедии. Английская версия базы знаний DBpedia в настоящее время описывает 6,0 млн сущностей, из которых 4,6 млн имеют аннотации, 1,53 млн имеют географические координаты и 1,6 млн изображений.

LCSH – предметные рубрики Библиотеки Конгресса США. Используется на международном уровне, часто в переводном варианте.

Из проектов, упомянутых в табл. 1, рассмотрим особенности портала DBNF (Data.bnf) Национальной библиотеки Франции как наиболее интересную реализацию решений, использующих связанные данные. Сайт основан на CubicWeb, платформе с открытым исходным кодом, которая позволяет извлекать данные из разных баз данных, созданных в разных форматах, и публиковать их в виде HTML-страниц, позволяющих отображать информацию пользователям через веб-браузер и в виде RDF для машинной обработки. Повторному использованию данных способствует применение онтологий семантической сети, включая FOAF для авторов, SKOS для концепций и RDA для основных элементов работ. Data.bnf согласован с виртуальным международным авторитетным файлом (VIAF), который представляет свои наборы данных в виде связанных открытых данных. Как отмечается в [6], библиотеки должны поощрять использование открытых ресурсов, подтверждая в то же время свою роль «гарантов качества», что является всё более острой потребностью в цифровом мире.

На некоторых порталах библиотек [3] предметные рубрики связаны со списком предметных рубрик для публичных библиотек (LEMB), опубликованном в формате с использованием пространства имён SKOS (Simple Knowledge Organization System), которые в свою очередь связаны с другими предметными списками, такими как LCSH (предметные рубрики Библиотеки Конгресса) или RAMEAU. Такого рода связи открывают пользователям дополнительные источники для поиска информации по интересующей их тематике. Особенностью модели SKOS является поддержка языковых тегов для лексических терминов. Тезаурусы, представленные в SKOS, потенциально являются многоязычными, что даёт дополнительные возможности их использования для построения семантических связей.

Использование связанных данных в проекте Единого открытого архива информации ГПНТБ России (ЕОАИ)

Как было отмечено ранее, применение связанных данных означает, с одной стороны, представление их в формате RDF для последующего использования и обработки информационными системами, а с

другой – обогащение информации, выдаваемой пользователям поисковых порталов. В данном исследовании мы рассматривали вопросы, относящиеся ко второму направлению, то есть обогащению информации, формируемой при выводе результатов поиска. В качестве базового набора данных, представленного в формате SKOS, мы используем многоязычный тезаурус GEMET [7].

В публикации [8] нами были приведены результаты исследования соответствия предметных рубрик ЕОАИ концептам тезауруса GEMET. В настоящее время таблица соответствия предметных рубрик с концептами GEMET насчитывает 1592 значения.

Одним из результатов работы по разработке информационной системы ЕОАИ явилось создание базы данных соответствия кодов обновлённой номенклатуры научных специальностей с индексами УДК [9]. Данная таблица, помимо применения в качестве средства дополнительной навигации на пользовательском портале ЕОАИ, позволила отобразить индексы УДК для проведения анализа возможности их использования для построения соответствия с концептами тезауруса GEMET.

Задачей проведённого исследования являлась оценка процентного соотношения записей в анализируемых электронных каталогах, для которых можно автоматически сформировать связи между индексами УДК и предметными рубриками с концептами тезауруса GEMET.

Целью исследования является использование сформированных связей для обогащения информации, отображаемой на портале ЕОАИ в результатах выполнения поисковых запросов пользователей.

Предметом исследования являются массивы библиографических записей электронного каталога ЕОАИ и базы данных «Экология: наука и технология». Библиографические записи электронного каталога ЕОАИ содержат описания документов с полными текстами, созданные в результате оцифровки печатных изданий из фондов ГПНТБ России, а также документы, опубликованные в журналах и научных изданиях ГПНТБ России. Библиографические записи БД «Экология: наука и технология» [10] анализируются, поскольку тезаурус GEMET представляет собой многоязычный тезаурус окружающей среды и интересно сопоставить, насколько процентные соотношения для полученных связей различаются для специализированного электронного каталога и электронного каталога общего назначения.

Методика исследования. Для формирования первоначальных массивов данных индексы УДК и предметные рубрики, содержащиеся в каждой библиографической записи, были выгружены из анализируемых баз данных САБ ИРБИС в построчный текстовый формат и загружены в таблицы СУБД MySQL. Аналогичная процедура была проведена с записями базы данных научных специальностей (БД ВАК), содержащей соответствия между номенклатурами научных специальностей и индексами УДК. Как было отмечено в [7], рубрикатор тезауруса GEMET был нами загружен в среду СУБД MySQL.

На втором этапе исследования была проведена нормализация таблиц индексов УДК и тематических рубрик с целью получения числа записей в электронных каталогах для каждого уникального значения индекса УДК и предметной рубрики. Далее в процессе программного анализа соответствия предметных рубрик и индексов УДК с концептами GEMET полученные данные были дополнены индексами найденных концептов GEMET.

Результаты исследования. В табл. 2 представлены результаты построения связей между библиографическими описаниями и концептами тезауруса GEMET, полученные в результате анализа записей электронного каталога ЕОАИ.

Таблица 2

Результаты построения связей между библиографическими описаниями и концептами тезауруса GEMET для электронного каталога ЕОАИ

Значение	Количество
Всего записей в базе данных	26 074
Записей, связанных с концептами GEMET на основе индекса УДК верхнего уровня	5507
Записей, для которых сформирована связь между УДК и страницей в DBpedia	5471
Уникальных страниц DBpedia (для связей с использованием УДК)	30
Записей, связанных с концептами GEMET на основе предметной рубрики	3450

Значение	Количество
Записей, для которых сформирована связь между предметной рубрикой и страницей в DBpedia	2312
Уникальных страниц DBpedia (для связей с использованием предметных рубрик)	327

В табл. 3 представлены результаты формирования связей, полученные в результате анализа и обработки библиографических записей электронного каталога БД «Экология: наука и технология».

Таблица 3

Результаты построения связей между библиографическими описаниями и концептами тезауруса GEMET для базы данных «Экология: наука и технология»

Значение	Количество
Всего записей в базе данных	52 559
Записей, связанных с концептами GEMET на основе индекса УДК верхнего уровня	10 952
Записей, для которых сформирована связь между УДК и страницей в DBpedia	10 761
Уникальных страниц DBpedia (для связей с использованием УДК)	26
Записей, связанных с концептами GEMET на основе предметной рубрики	16 769
Записей, для которых сформирована связь между предметной рубрикой и страницей в DBpedia	10 304
Уникальных страниц DBpedia (для связей с использованием предметных рубрик)	223

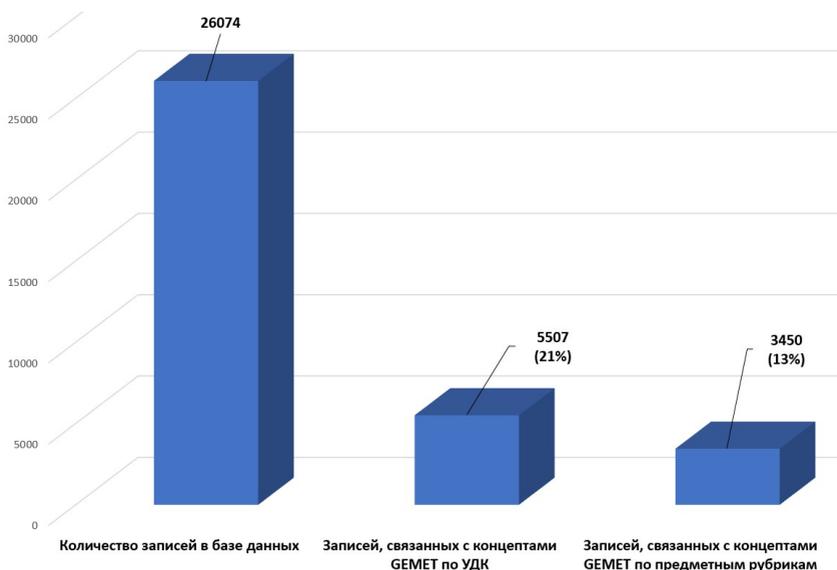


Рис. 1. Процентное соотношение записей, связанных с концептами GEMET в электронном каталоге ЕОАИ

На рис. 1 и 2 представлены процентные соотношения записей, для которых были сформированы связи с концептами GEMET в анализируемых базах данных. В обоих случаях связи на основе УДК были сформированы примерно для 20% записей. Но, как и следовало ожидать, связей на основе предметных рубрик значительно больше в БД «Экология: наука и технология», поскольку её тематика точнее соответствует концептам GEMET.

Для пользователей портала ЕОАИ интерес представляют не формально построенные связи, а информация, содержащаяся в определениях терминов тезауруса и связи со страницами в DBpedia. На рис. 3 приведён пример, показывающий результат расширения объёма информации, выдаваемой пользователю на основе связанных данных.

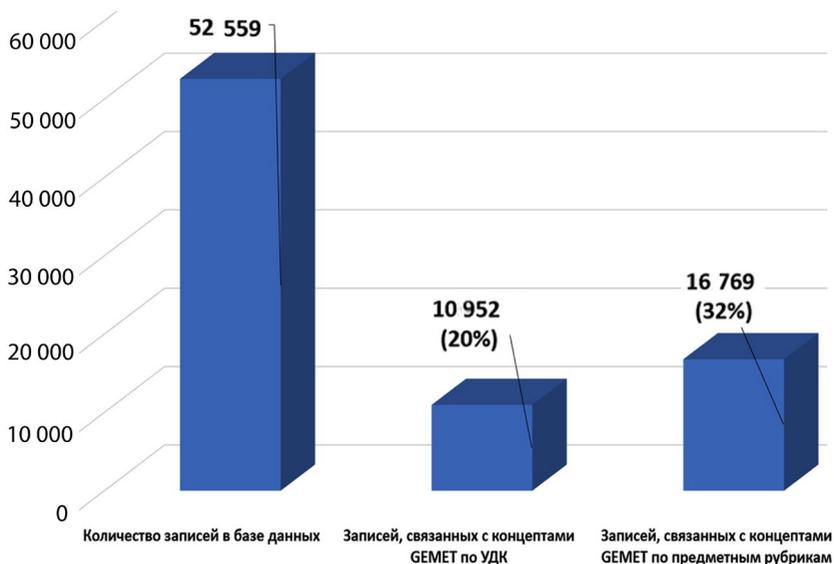


Рис. 2. Процентное соотношение записей, связанных с концептами GEMET в базе данных «Экология: наука и технология»

Как следует из данных, приведённых в табл. 1, связей с уникальными страницами DBpedia, построенных по индексам УДК, сравнительно немного. С целью расширения возможностей использования связанных данных на портале ЕОАИ эти связи были дополнены ссылками на страницы российской онлайн-энциклопедии РУВИКИ. В табл. 4 приведён пример сформированных связей.

Библиографическое описание:

Бугаец, Евгений Андреевич. Теория стереофотограмметрического воздушного и наземного методов создания топографических карт: лекция 2-я : для студентов горного факультета специальности "Маркшейдерское дело" / Е. А. Бугаец; Всесоюзный заочный политехнический институт, Кафедра "Маркшейдерское дело и геодезия". - Москва : [б. и.], 1960. - 65, [2] с. : ил. Рубрики: Фотограмметрия, Топографические карты -- Составление и редактирование

УДК: 528.721.2

Код УДК	Индекс концепта GEMET	Метка концепта GEMET на русском языке	Определение концепта GEMET на русском языке
528	6204	фотограмметрия	Процесс осуществления измерений с применением фотографий; часто используется при составлении карт на основе аэрофотосъемки, а также для военной разведки, медицинских и промышленных исследований и пр.
	3638	геодезия	Отрасль геофизики, включающая определение размеров и формы Земли, гравитационного поля Земли, положение точек в земной коре в системе координат, привязанной к Земле.

Метка концепта GEMET на русском языке	URL страницы в DBpedia
фотограмметрия	http://dbpedia.org/resource/Photogrammetry
геодезия	http://dbpedia.org/resource/Geodesy

Рис. 3. Информация, получаемая из базы данных на основе связи между индексом УДК и концептами GEMET

**Пример дополнения сформированных связей ссылками
на страницы в РУВИКИ**

Индекс УДК	Число записей в БД	Индекс концепта GEMET	Связанная страница DBpedia	Страница в РУВИКИ
53	111	6230	http://dbpedia.org/resource/Physical_science	https://ru.ruwiki.ru/wiki/Физические_науки
553	431	14 837	http://dbpedia.org/resource/Mining	https://ru.ruwiki.ru/wiki/Горное_дело
626	154	4085	http://dbpedia.org/resource/Hydraulics	https://ru.ruwiki.ru/wiki/Гидравлика
697	153	3894	http://dbpedia.org/resource/Heater	https://ru.ruwiki.ru/wiki/Отопление_вентиляция_и_кондиционирование
631	355	11 152	http://dbpedia.org/resource/Agronomy	https://ru.ruwiki.ru/wiki/Агрономия
656	335	8653	http://dbpedia.org/resource/Transport_network	https://ru.ruwiki.ru/wiki/Транспортная_инфраструктура

Заключение

Широкое внедрение технологии связанных данных на поисковых порталах современных библиотек является трудоёмкой задачей. Как показывает опыт зарубежных библиотек, только крупные библиотеки могут полностью преобразовать огромный объём данных в связанные данные. В связи с этим для массовых библиотек могут представлять интерес технологии, позволяющие построить связи в автоматизированном режиме, с использованием возможностей словарей RDF, доступных в формате модели SKOS. В свою очередь, данные этих словарей можно расширять, как это показано нами в настоящей статье на примере построения связей между концептами GEMET и страницами в РУВИКИ – в дополнение к существующим связям со страницами в DBpedia. Отдельный интерес представляет использование связанных данных имён авторов, представленных в международном авторитет-

ном файле VIAF, но в настоящее время интерактивный доступ к его ресурсам с использованием API закрыт для российских пользователей. Имеется возможность скачивания текущей версии VIAF для использования в качестве локального ресурса, но это требует значительных объёмов дискового пространства и не является актуальным для использования в проекте ЕОАИ.

Список источников

1. **Hallo M. et al.** Current state of Linked Data in digital libraries // Journal of Information Science. 2016. Т. 42. № 2. С. 117–127.
2. **Alvite-Diez M. L.** Linked open data portals: functionalities and user experience in semantic catalogues // Online Information Review. 2021. Т. 45. № 5. С. 946–963.
3. **Wang Y., Yang S. Q.** Linked data technologies and what libraries have accomplished so far // International Journal of Librarianship. 2018. Т. 3. № 1. С. 3–20.
4. **Zapounidou S., Sfakakis M., Papatheodorou C.** Representing and integrating bibliographic information into the Semantic Web: A comparison of four conceptual models // Journal of information science. 2017. Т. 43. № 4. С. 525–553.
5. **The Linked** Open Data Cloud. URL: <https://lod-cloud.net/> (дата обращения: 12.09.2024).
6. **Grimaldi E.** The evolution of Data. bnf. fr: past, present and future of the BnF linked open data project // J LIS. it. 2024. Т. 15. № 2. С. 119–133.
7. **General** Multilingual Environmental Thesaurus (GEMET). URL: <https://www.eionet.europa.eu/gemet/en/themes/> (дата обращения: 12.09.2024).
8. **Гончаров М. В., Колосов К. А.** Формирование расширенных метаданных в информационной системе ЕОАИ ГПНТБ // Научные и технические библиотеки. 2023. № 11. С. 84–98.
9. **Зайцева Е. М.** Развитие лингвистических средств тематического поиска в библиотечно-информационных системах // Научные и технические библиотеки. 2023. № 11. С. 66–83.
10. **Бычкова Е. Ф., Колосов К. А.** Анализ возможностей автоматического реферирования статей на примере источников базы данных «Экология: наука и технологии» ГПНТБ России // Научные и технические библиотеки. 2023. № 10. С. 99–120.

References

1. **Hallo M. et al.** Current state of Linked Data in digital libraries // Journal of Information Science. 2016. T. 42. № 2. C. 117–127.
2. **Alvite-Diez M. L.** Linked open data portals: functionalities and user experience in semantic catalogues // Online Information Review. 2021. T. 45. № 5. C. 946–963.
3. **Wang Y., Yang S. Q.** Linked data technologies and what libraries have accomplished so far // International Journal of Librarianship. 2018. T. 3. № 1. C. 3–20.
4. **Zapounidou S., Sfakakis M., Papatheodorou C.** Representing and integrating bibliographic information into the Semantic Web: A comparison of four conceptual models // Journal of information science. 2017. T. 43. № 4. C. 525–553.
5. **The Linked** Open Data Cloud. URL: <https://lod-cloud.net/> (Accessed: 12.09.2024).
6. **Grimaldi E.** The evolution of Data. bnf. fr: past, present and future of the BnF linked open data project // JILIS. it. 2024. T. 15. № 2. C. 119–133.
7. **General** Multilingual Environmental Thesaurus (GEMET). URL: <https://www.eionet.europa.eu/gemet/en/themes/> (Accessed: 12.09.2024).
8. **Goncharov M. V., Kolosov K. A.** Formirovanie rasshirenykh metadannykh v informatsionnoi sisteme EOAI GPNTB // Nauchny'e i tekhnicheskie biblioteki. 2023. № 11. S. 84–98.
9. **Zai'tceva E. M.** Razvitie lingvisticheskikh sredstv tematicheskogo poiska v bibliotekho-informatsionnykh sistemakh // Nauchny'e i tekhnicheskie biblioteki. 2023. № 11. S. 66–83.
10. **By'chkova E. F., Kolosov K. A.** Analiz vozmozhnostei avtomaticheskogo referirovaniia statei na primere istochnikov bazy dannykh «E'kologiya: nauka i tekhnologii» GPNTB Rossii // Nauchny'e i tekhnicheskie biblioteki. 2023. № 10. S. 99–120.

Информация об авторах / Authors

Гончаров Михаил Владимирович – канд. техн. наук, доцент, ведущий научный сотрудник, руководитель группы перспективных исследований и аналитического прогнозирования ГПНТБ России; доцент Московского государственного лингвистического университета, Москва, Российская Федерация
goncharov@gpntb.ru

Mikhail V. Goncharov – Cand. Sc. (Engineering), Associate Professor, Leading Researcher, Head, Group for Perspective Research and Analytic Forecasting, Russian National Public Library for Science and Technology; Associate Professor, Moscow State Linguistic University, Moscow, Russian Federation
goncharov@gpntb.ru

Колосов Кирил Анатольевич –
канд. техн. наук, ведущий научный
сотрудник ГПНТБ России; доцент
Московского государственного
лингвистического университета,
Москва, Российская Федерация
kolosov@gpntb.ru

Kirill A. Kolosov – Cand. Sc.
(Engineering), Leading Researcher,
Russian National Public Library for
Science and Technology; Associate
Professor, Moscow State Linguistic
University, Moscow,
Russian Federation
kolosov@gpntb.ru