

**Связывание системы Топонимия Карелии  
с другими топонимическими базами данных**

**А. А. Крижановский**

*Институт прикладных математических исследований  
Карельского научного центра РАН,  
г. Петрозаводск, Российская Федерация,  
andrew.krizhanovsky@gmail.com*

**АННОТАЦИЯ**

**Введение.** В Институте языка, литературы и истории Карельского научного центра РАН собирают картотеку топонимов Карелии и сопредельных областей с 1970 г. Собранный картотека содержит более 300 тыс. топонимов на русском, карельском и вепском языках. Более 20 лет идёт работа по переводу бумажной картотеки в электронный формат и добавлению новых материалов полевых исследований. Создана геоинформационная онлайн-система Топонимия Карелии (ТопКар). В связи с ростом числа топонимических ресурсов встаёт задача унификации данных и связывания их в единую сеть. Всё чаще в различных областях (библиотечное дело, биоинформатика) таким объединяющим звеном выступает платформа Викиданных. На примере 11 топонимических баз данных, включая ТопКар, показано их объединение в единую сеть. Проанализировано то, какими особенностями должны обладать топонимические базы для их включения в сеть связанных данных.

**Цель:** представить современные топонимические базы, выявить их особенности, обеспечивающие включение в существующую единую сеть связанных данных.

**Материалы исследования:** современные онлайн топонимические базы данных.

**Результаты и научная новизна.** Описан опыт разработки ресурса Топонимия Карелии с точки зрения интеграции с другими онлайн-ресурсами. Объяснены трудности и перечислены ключевые моменты, дающие возможность включать топонимические данные в сеть связанных данных. Представлен обзор современных топонимических баз данных всего мира (GeoNames, GNIS, GNS, TGN) и региональных (Топонимия Карелии, Composite Gazetteer of Antarctica, Топонимы норвежских полярных территорий, Järviwiki, Топонимическая база данных Латвии, База нижнелужицких топонимов и Monasticon Hibernicum). Научная новизна заключается в том, что выявленные особенности этих ресурсов могут служить рекомендациями при разработке новых топонимических систем для успешного включения в сеть связанных данных.

**Ключевые слова:** топонимическая база данных, топонимия, Викиданные, связанные данные, SPARQL-запрос, GeoNames, GNIS, GNS, TGN

*Для цитирования:* Крижановский А. А. Связывание системы Топонимия Карелии с другими топонимическими базами данных // Вестник угроведения. 2025. Т. 15. № 3 (62). С. 578–587.

**Linking the Toponymy of Karelia system  
with other toponymic databases**

**A. A. Krizhanovsky**

*Institute of Applied Mathematical Research,  
Karelian Research Center of the Russian Academy of Sciences,  
Petrozavodsk, Russian Federation,  
andrew.krizhanovsky@gmail.com*

**ABSTRACT**

**Introduction:** researchers of the Karelian Research Centre of the Russian Academy of Sciences have been compiling a card index of toponyms from Karelia and adjacent regions since 1970. This index contains over 300,000 toponyms in the Russian, Karelian, and Vepsian languages. The process of digitizing the paper-based card index and adding new field research materials has been ongoing for more than 20 years. The Toponymy of Karelia (TopKar) geoinformation online system has been developed. The number of toponymic resources is growing, making it necessary to standardize data and connect them into a unified network. In various fields (such as library science and bioinformatics), the Wikidata database serves as a unifying link. Let's demonstrate this network integration using 11 toponymic databases, including TopKar. Let's identify the unique features of toponymic databases to enable their inclusion in the linked data.

**Objective:** to present modern toponymic databases; to identify their features that ensure inclusion in the existing unified network of the linked data.

**Research materials:** modern online toponymic databases.

**Results and novelty of the research:** the experience of developing the Toponymy of Karelia system and linking it with other online resources is described. The difficulties are explained, and key points are outlined that facilitate the inclusion of toponymic databases into the linked data. Global toponymic databases (GeoNames, GNIS, GNS, TGN) and regional databases (TopKar, Composite Gazetteer of Antarctica, Place names in Norwegian polar areas, Järviwiki, Latvian toponymic names database, Lower Sorbian place name, and Monasticon Hibernicum) are reviewed.

New developing toponymic databases will be more successfully included in the linked data if the identified features of databases are considered.

**Key words:** toponymic database, toponymy, Wikidata, linked data, SPARQL queries, GeoNames, GNIS, GNS, TGN

*For citation:* Krizhanovsky A. A. Linking the Toponymy of Karelia system with other toponymic databases // Vestnik ugrovedenia = Bulletin of Ugric Studies. 2025; 15 (3/62): 578–587.

## Введение

Организация и публикация топонимических данных в виде, пригодном к компьютерной обработке, становятся всё более актуальными. Однако обзорные работы по топонимическим базам заметно устарели [2; 13]. В исследовании F. Ormeling [13] представлен систематический обзор крупных международных топонимических баз данных, в том числе: GeoNames, GNIS, GNS, TGN. Эти ресурсы широко используются в лингвистических, историко-культурных и картографических исследованиях.

Интеграция подобных баз остаётся сложной задачей. Различия в структуре, терминологии, степени детализации и форматах хранения препятствуют их сопоставлению и совместному использованию. Решение этой проблемы предлагается через применение технологий связанных данных (Linked Data) и онтологически согласованных форматов описания.

В последние годы особое значение приобретают Викиданные (Wikidata) — граф знаний, сообщество, онлайн-платформа и проект Викимедиа [18], который становится важным хранилищем и точкой соприкосновения для множества лексикографических, географических и краеведческих проектов. Это крупнейший проект связанных открытых данных [8]. Хотя в Викиданных уже представлены многие крупные топонимы, включая административные единицы и географические объекты, системная интеграция региональных топонимических баз остаётся редкой и слабо документированной. Существуют широкие возможности для создания методик и прототипов интеграции, способных расширить потенциал таких ресурсов и повысить их взаимосвязь с другими источниками открытых данных.

Успешная интеграция информационных ресурсов требует построения системы глубоких ссылок (deep links), связывающих внутренние страницы внешних источников [14]. Викиданные в этом случае выступают в роли универ-

сального посредника, объединяющего топонимические ресурсы с помощью внешних идентификаторов [17].

До появления Викиданных не было общепринятого идентификатора (подобного DOI для научных статей) для однозначного определения топонима. Эту функцию выполняет QID — уникальный идентификатор объекта в Викиданных [14]. Благодаря ссылкам на QID (аналогично ссылкам на постоянные и неизменные DOI в статьях) топонимические базы данных могут быть объединены в единую сеть.

В этой статье решим следующие задачи: (1) рассмотрим региональные и международные топонимические ресурсы (TopKar, GeoNames, GNIS и др.); (2) выявим особенности, позволившие включить их в Викиданные; (3) опишем преимущества, которые даёт объединение Викиданных и локальных топонимических баз данных для разработчиков, исследователей и пользователей.

## Материалы и методы

Материалом исследования являются топонимические базы данных, доступные онлайн и включённые в Викиданные с помощью внешнего идентификатора (указан далее для каждой базы).

В работе использованы следующие методы:

1. расширенный SPARQL-анализ [4] — формирование SPARQL-запросов для отбора баз данных, связанных с топонимами через внешние идентификаторы и обеспечивающих глубокую интеграцию (deep linking);

2. сравнительный анализ и систематизация — сопоставление структуры и объёма связей каждой базы, а также обзор литературы по топонимическим веб-ресурсам [2; 8; 13];

3. экспертная оценка и контент-анализ — изучение форматов и моделей представления топонимических данных; ручная проверка наличия и корректности внешних идентификаторов и степени детализации ссылок между базами.

## Результаты

Опишем принцип отбора топонимических баз данных для исследования и внешний идентификатор в Викиданных. Затем рассмотрим 11 топонимических баз данных и выявим те особенности, благодаря которым эти базы были включены в единую сеть связанных данных.

### 1. Выбор топонимических баз и внешний идентификатор в Викиданных

Существует огромное число баз данных, работающих онлайн (около 9 тыс. наиболее известных, см. <https://w.wiki/DZq9>). Число топонимических баз данных в мире неизвестно. Перед нами стояла задача отбора материала. Нужно выбрать такие топонимические базы данных, которые связаны гиперссылками с Викиданными. Причём должно быть связывание не главных страниц проектов (одна взаимная ссылка), а связывание на уровне объектов, когда ссылками связывается один и тот же объект, описанный в разных базах данных. То есть топонимическая база данных имеет глубокое связывание с Викиданными [14].

В Викиданных объект значим (имеет право на создание и существование), если выполняется хотя бы один из трёх критериев: (1) содержит рабочую ссылку на страницу одного из проектов Викимедиа; (2) представляет собой экземпляр чётко определяемой концептуальной или материальной сущности; (3) необходим для структурной целостности базы данных (например, служит связующим элементом между другими объектами Викиданных).

Каждый объект однозначно определяется идентификатором QID, уникальным числом с префиксом Q. Идентификатор QID, аналогично DOI, однозначно определяет веб-страницу объекта Викиданных в Интернете. Данные в Викиданные добавляются в виде утверждений, а именно в виде троек: объект – свойство – значение. Объекты безусловно связаны друг с другом с помощью утверждений. На странице объекта на сайте Викиданных в разделе «Утверждения» можно добавить связующую ссылку («Свойство»), нажав кнопку «Редактировать». Каждое свойство имеет уникальный идентификатор, число с префиксом P (от слова Property) [8].

Для поиска в Викиданных геоинформационных систем, связанных с топонимией, мы использовали несколько способов: (1) стандартная страница поиска Викиданных, при указании поиска только по свойствам; (2) поиск

с помощью SPARQL-запроса (<https://w.wiki/DRKf>) таких объектов, у которых указано, что “main subject” = “toponym” (то есть главной темой объекта являются топонимы); (3) поиск по более сложному запросу (<https://w.wiki/Da4S>) таких онлайн баз данных, которые связаны с топонимами.

При этом у такой топонимической системы должно быть «своё» свойство в Викиданных с внешним идентификатором. Наличие «своего» свойства в виде внешнего идентификатора указывает, что ведётся достаточно планомерная работа по включению ссылок на топонимические объекты. Дальше на конкретных примерах мы покажем, что это за свойство «внешний идентификатор». Это ограничение существенно сузило поиск топонимических баз данных, рассмотренных далее.

### 2. Топонимия Карелии и Викиданные

Институт языка, литературы и истории Карельского научного центра РАН (ИЯЛИ КарНЦ РАН) существует с 1930 г. На протяжении последних пяти десятилетий сотрудники института проводят полевые экспедиции, собирая топонимы Карелии и соседних областей. Топонимы в этой (бумажной) картотеке привязаны к топографическим картам. Картотека содержит более 300 тыс. топонимов на русском, карельском и вепсском языках [1].

В начале 2000-х гг. под руководством И. И. Муллонен в ИЯЛИ КарНЦ РАН началась оцифровка картотеки по топонимии Карелии. Была создана локальная база данных и программа, работавшая без доступа к Интернету. Этой системой топонимисты пользовались около 20 лет; в ней собрано более 60 тыс. топонимов.

В 2022 г. в Карельском научном центре РАН начали разрабатывать геоинформационную онлайн-систему Топонимия Карелии (ТопКар), доступную в сети Интернет (<http://topkar.krc.karelia.ru>). Работы были начаты благодаря гранту РФ № 22-28-00362 (“Создание открытой геоинформационной системы по топонимии Карелии на примере Арктической зоны республики”). В эту геоинформационную систему были включены данные разработанной ранее локальной ГИС, также исследователи добавляют в неё топонимы из бумажной картотеки и материалы новых экспедиций.

Опишем свой опыт разработки ресурса Топонимия Карелии (ТопКар), который связан

гиперссылками с объектами Викиданных. На рис. 1 показана суть идеи связывания научной базы данных, Википедии и Викиданных, и пронумерованы три следующих подраздела, поясняющих подход на примере базы данных Топонимия Карелии.

**Ссылка первая: с объекта Топонимия Карелии на объект Викиданных**

На рис. 2 представлена страница топонима Lahnaoja в разработанном веб-ресурсе Топонимия Карелии. Обратите внимание на два идентификатора в левой верхней части рисунка, а именно:

1) Идентификатор Викиданных Q36555214, подчёркивание указывает, что текст идентификатора является гиперссылкой на страницу Викиданных (<https://www.wikidata.org/wiki/Q36555214>).

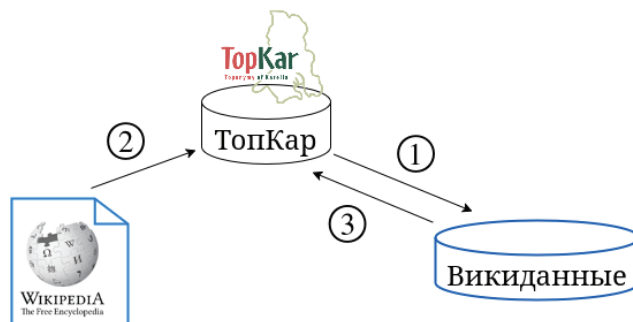
2) Идентификатор топонима в системе Топонимия Карелии (здесь *TopKar ID* равен 26007). Этот идентификатор однозначно определяет URL этой страницы: <http://topkar.krc.karelia.ru/en/dict/toponyms/26007>. Эта возможность получения URL страницы по идентификатору

топонима в базе данных важна на шаге 3 для связки Викиданных с TopKar.

Если выбрать в системе TopKar пункт меню «Топонимы / С Викиданными», то откроется страница со списком всех топонимов, имеющих ссылку на соответствующий объект Викиданных.

При заполнении в TopKar ссылки на объект Викиданных вся сложность заключается в том, чтобы найти этот соответствующий объект. Его может и не быть, поскольку Викиданные и TopKar основаны на разных принципах отбора материала. По-видимому, большое количество топонимов TopKar отсутствуют в Викиданных. Например, отсутствуют такие виды объектов, как: кладбище, дом, дорога, крест, канава. Объекты этих видов можно увидеть по ссылке [http://topkar.krc.karelia.ru/misc/geotypes?search\\_name=&sort\\_by=name\\_ru&portion=122](http://topkar.krc.karelia.ru/misc/geotypes?search_name=&sort_by=name_ru&portion=122), нужно кликнуть по числу объектов в соответствующей строке.

Таким образом, ссылка с топонима TopKar на топоним Викиданных задана и представлена на рис. 2.



**Рис. 1.** Взаимосвязь топонимической системы TopKar, Викиданных и Википедии

The screenshot shows the web interface for the toponym 'Lahnaoja'. At the top left, the name 'Lahnaoja' is displayed. Below it, the Wikidata ID 'Q36555214' and TopKar ID '26007' are shown. To the right of the name are icons for 'Return to list', 'Edit', 'Delete', 'Create another one', and 'In this settlement'. Below the name is a map of the region with a blue location pin. To the right of the map is a list of metadata: 'Other names: Lahnoja, Лагноя', 'Wrong names:', 'Geotype: stream, brook', 'Region, district, settlement: Republic of Karelia, Kondopozhsky District', 'Region, district, selsovet, settlement (early 20th century): Republic of Karelia, Petrovsky District, Munozersky, Munozero, Pogost', 'Locative form:', 'Ethnos territory: Ludics', 'Language of origin:', 'Etymology:', and 'Main information: Ручей недалеко от д. Мунозеро, впадает в р. Пяла. (A stream not far from the village of Munozero, flows into the Pyala River.)'. A scale bar at the bottom left of the map shows 300 m and 1000 ft.

**Рис. 2.** Топоним Lahnaoja в системе Топонимия Карелии с заполненным значением идентификатора Викиданных

### Ссылка вторая: со статьи Википедии на объект в системе TopKar

Топонимические статьи TopKar содержат информацию, которая может использоваться в Википедии. По правилам Википедии (Wikipedia:Citing sources) при добавлении какой-либо информации необходимо указывать ссылку на источник. Для удобства и единообразия ссылок была создана страница «Шаблон:Topkar» (<https://ru.wikipedia.org/wiki/Шаблон:Topkar>).

Теперь для формирования развёрнутой (библиографической) ссылки на топоним TopKar в тексте статьи Википедии достаточно написать `{{Topkar|Lahnaoja|26007|ref=TopKar}}`, где *Lahnaoja* – название топонима, а 26007 – идентификатор топонима в TopKar. Код шаблона разворачивается в надпись: «Lahnaoja // Топонимия Карелии. topkar.krc.karelia.ru. ИЯЛИ КарНЦ РАН» с гиперссылкой на топоним номер 26007 в TopKar.

Вот пример более простого шаблона в Английской Википедии, связанного с топонимией Латвии: <https://en.wikipedia.org/wiki/Template:Vietvardi.lgia.gov.lv>. Этот шаблон (без параметров) генерирует неизменяемый текст с указанием источника. Далее мы рассмотрим эту Топонимическую базу данных Латвии.

### Ссылка третья: с объекта Викиданных на объект TopKar

Основной единицей Викиданных является объект (например, объект *Lagnoya* – ручей в Карелии, см. <https://www.wikidata.org/wiki/Q36555214>). Основной единицей описания объекта в Викиданных является свойство. Свойства делятся на две группы: (1) утверждения (Statements), например, «Lagnoya – это экземпляр ручья» или «Lagnoya имеет длину 15 км»; (2) идентификаторы (Identifiers) – указатели, связывающие данный объект с конкретными записями внешних (по отношению к Викиданным) баз данных.

В рассматриваемом объекте *Lagnoya* (Q36555214) указано свойство «GeoNames ID». То есть база данных GeoNames имеет в Викиданных свойство GeoNames ID (<https://www.wikidata.org/wiki/Property:P1566>) для указания на страницах объектов Викиданных соответствующих идентификаторов в базе GeoNames.

База данных TopKar является внешней по отношению к Викиданным. Поэтому свойство, связывающее объект Викиданным с TopКаром

(*TopKar ID*), является внешним идентификатором для Викиданных.

Теперь конкретизирован описанный выше критерий отбора топонимических баз данных: наличие внешнего идентификатора для рассматриваемой базы. Поскольку этот критерий для нас ключевой, в дальнейшем при описании баз данных мы будем указывать соответствующие внешние идентификаторы.

### 3. GeoNames

GeoNames – это географическая база данных онлайн ([www.geonames.org](http://www.geonames.org)), на которую завязаны многие топонимические базы [19]. Внешний идентификатор Викиданных *Geonames ID* (P1566) используется для связи с базой данных GeoNames, которая содержит информацию о географических объектах. Из 25 млн топонимов GeoNames 4 млн уже включены в Викиданные с помощью этого идентификатора.

В Английской Википедии есть специальный шаблон `{{Geonames link}}`, который позволяет редакторам удобно связывать тексты статей энциклопедии с географической базой данных GeoNames. Двойные фигурные скобки приняты для обозначения шаблонов в вики-проектах. В этом шаблоне предусмотрено два параметра: идентификатор (для генерации гиперссылки на объект в базе GeoNames) и название объекта.

### 4. Geographic Names Information System

Информационная система географических названий (Geographic Names Information System, кратко GNIS) – это географическая база данных топонимов на территории США и прилегающих территориях. База GNIS разработана Геологической службой США в сотрудничестве с Советом США по географическим названиям.

Внешний идентификатор Викиданных *GNIS Feature ID* (P590) используется для связи с базой данных GNIS. Идентификатор используется 1,5 млн раз, всего GNIS содержит 2,3 млн топонимов. Шаблон `{{GNIS}}` для генерации ссылок в Английской Википедии используется в 58 тыс. статей, в Русской Википедии – в 1300 статьях.

База данных GNIS регулярно публикуется в виде дампов данных в формате CSV. Каждая запись базы включает тип географического объекта, его координаты, город, округ и штат, к которому он относится, высоту над уровнем моря, дату создания записи, исходный картографический источник, альтернативные названия, источники (исторические, законодательные) [9].

## 5. GEOnet Names Server

GEOnet Names Server (кратко GNS) – это онлайн-база географических названий ([geonames.nga.mil](http://geonames.nga.mil)), управляемая Национальным агентством геопроостранственной разведки США в сотрудничестве с Советом США по географическим названиям. База GNS хранит топонимы тех мест, которые находятся за пределами США [13].

База GEOnet является основным источником информации для многих других топонимических баз. Карты в системе GEOnet отображаются с помощью сервисов Open Geospatial Consortium, Google Карты и MapQuest [13].

Внешний идентификатор Викиданных *GNS Unique Feature ID* (P2326) связывает около 3,6 млн топонимов Викиданных с базой данных GNS, содержащей 14,7 млн топонимов. В статьях Английской Википедии для ссылок на GNS используется шаблон `{{GEOnet3}}`, который применяется примерно 43 тыс. раз. Данные GNS распространяются под максимально открытой лицензией – они находятся в общественном достоянии (Public Domain).

## 6. Getty Thesaurus of Geographic Names

Getty Thesaurus of Geographic Names (кратко TGN) – это тезаурус топонимов Исследовательского института Гетти в Лос-Анджелесе. В 2022 г. тезаурус TGN содержал около 3 млн записей. Каждая запись включает название объекта, его тип, географические координаты и текстовое описание. Названия мест могут включать варианты названий, экзонимы и исторические названия. В базе представлены различные варианты названий, включая экзонимы и исторические наименования, а также указаны временные рамки их использования. Указаны источники данных. Тезаурус TGN охватывает широкий временной диапазон – от доисторических времён до настоящего времени, а географически топонимы покрывают карту всего мира [13].

В Викиданных внешний идентификатор *Getty Thesaurus of Geographic Names ID* (P1667) встречается примерно в 90 тыс. топонимах. Проект TGN предоставляет доступ к данным через поисковую форму на сайте [www.getty.edu](http://www.getty.edu), а также позволяет скачать базу данных в различных форматах. Данные распространяются по открытой лицензии ODC-By 1.0.

## 7. Composite Gazetteer of Antarctica

Примером международного сотрудничества является топонимическая база данных Composite Gazetteer of Antarctica (CGA),

работающая под эгидой Научного комитета по изучению Антарктики (SCAR). Соответствующее свойство Викиданных называется *SCAR Composite Gazetteer place ID* (P3230). Почти все антарктические топонимы CGA (около 20 тыс.) уже включены с помощью ссылок в Викиданные.

Примером международного языкового конструирования являются как сами антарктические топонимы, придуманные путешественниками и исследователями разных стран за последние 250 лет, так и географические термины, заимствованные из других языков или специально придуманные для описания местного рельефа [10].

## 8. Топонимы норвежских полярных территорий

Норвежский полярный институт работает над базой топонимов *Place names in Norwegian polar areas* (<https://placenames.npolar.no>). Институт опубликовал набор данных с открытой лицензией в формате JSON [12], содержащий 23,5 тыс. описаний полярных мест. Как указано в заявке (<https://www.wikidata.org/?curid=55187010>) по созданию соответствующего свойства *Norwegian Polar Institute place name ID* (P5391) на Викиданных, именно этот набор данных использовался для автоматической загрузки в Викиданные. При внесении и обработки большого объёма однотипных данных в Викиданных используется механизм ботов [4].

## 9. Järviwiki

В 2011 г. Финский институт окружающей среды (SYKE) запустил сайт финских озёр Järviwiki ([www.jarviwiki.fi](http://www.jarviwiki.fi)) для гражданско-го мониторинга и пополнения базы данных. Järviwiki содержит сведения о температуре воды, состоянии поверхности, прозрачности и цветении водорослей, а также позволяет жителям Финляндии сохранять информацию о своих озёрах [7]. Такие сведения, собранные добровольцами, называют Volunteered Geographic Information или VGI. Информацию о водоёмах редакторы могут дополнять связанными с ними легендами и народными верованиями [11].

В Викиданных 18 тыс. топонимов связаны с озёрами Järviwiki через внешний идентификатор *Finnish Lake ID* (P3394). В базе Järviwiki всего 56 тыс. водоёмов.

## 10. Топонимическая база данных Латвии

В 1998 г. была создана База данных географических названий Латвии [5]. Латвийское

агентство геопространственной информации ведёт сайт топонимической базы данных Латвии (<https://vietvardi.lgia.gov.lv>), содержащей около 132 тыс. объектов. К сожалению, интерфейс сайта представлен только на латышском языке.

В 2019 г. Национальная библиотека Латвии начала пополнять записи предметных рубрик (authority records) внешними ссылками на другие связанные источники данных, в том числе ссылками на эту базу данных названий мест Латвии (11 % от всех ссылок в библиотеке) и на Викиданные (5,3 %) [6].

В Викиданных можно ссылаться на базу латвийских топонимов благодаря наличию свойства *Latvian toponymic names database ID* (P2496). Викиданные содержат 9674 ссылок на объекты Топонимической базы данных Латвии, что составляет около 7 % от всех топонимов в базе.

### **11. База нижнелужицких топонимов**

Серболужицкий институт в Германии разрабатывает двуязычный веб-сайт [dolnosorbiski.de/niedersorbisch.de](http://dolnosorbiski.de/niedersorbisch.de), посвящённый различным аспектам нижнелужицкого языка.

Собственные имена были извлечены из словарей. Собранный материал был систематизирован, унифицирован, дополнен и опубликован в гипертекстовой форме с обширными ссылками на дополнительные источники. Основная цель сервиса – облегчить поиск правильных форм собственных имён. База данных уже используется для распознавания именованных сущностей (Named Entities) в нижнелужицких текстах [16].

База данных нижнелужицкого языка содержит названия более 1000 населённых пунктов в Бранденбурге и более 100 в Саксонии, а также свыше 4 тыс. ссылок на внешние базы данных, такие как Geonames, OpenStreetMap и Викиданные. В свою очередь, в Викиданных 902 объекта имеют свойство *Lower Sorbian place name ID* (P8661), связывающее их с нижнелужицкими топонимами на сайте [dolnosorbiski.de](http://dolnosorbiski.de).

### **12. Monasticon Hibernicum**

Monasticon Hibernicum – это монастырский реестр Ирландии, включающий данные о раннехристианских церквях, соборах, монастырях, обителях и скитах Ирландии, о которых сохранились исторические, археологические или топонимические сведения [15]. Полное название реестра «Early Christian Ecclesiastical Settlement in Ireland 5th to 12th Centuries»

(Раннехристианские церковные поселения в Ирландии с V по XII вв.). Эта «энциклопедия в форме компьютерной базы данных с широкими возможностями поиска», как пишут авторы сайта ([monasticon.celt.dias.ie](http://monasticon.celt.dias.ie)), была подготовлена для запуска в Интернете при помощи Дублинского института перспективных исследований.

В Викиданных для ссылки на этот ресурс используется свойство *Monasticon Hibernicum database ID* (P12272), однако на данный момент оно применено всего в 10 объектах. Всего же сайт содержит около 5,5 тыс. ирландских топонимов.

### **13. Общие свойства топонимических баз данных**

Рассмотренные топонимические системы имеют ряд общих свойств. Базы данных доступны онлайн без какой-либо обязательной регистрации пользователей. У каждого топонима есть своя отдельная страница. У большинства ресурсов данные опубликованы по открытой лицензии.

Итак, у каждого топонима (объекта) в Викиданных есть уникальный идентификатор QID, аналогичный уникальному и постоянному DOI у статей.

Из 11 только три топонимические базы (GeoNames, База нижнелужицких топонимов, TopKap) содержат ссылки с топонимов на объединяющий их QID Викиданных. И наоборот, в Викиданных есть ссылки из объекта с идентификатором QID на соответствующие топонимы всех рассмотренных баз данных.

Отметим следующую особенность. Не обязательно топонимический ресурс имеет огромное число ссылок в Викиданных. Например, монастырский реестр Ирландии имеет всего 10 ссылок, но сама база Monasticon Hibernicum содержит около 5,5 тыс. ирландских топонимов. То есть работа по включению ссылок и связыванию ресурсов для некоторых проектов ещё только началась.

### **14. Проблемы с включением топонимических баз данных**

Можно указать отечественную базу данных «Топонимия Костромской области» доступную онлайн ([lexrus.ru](http://lexrus.ru)) [3]. К сожалению, эта база не соответствует концепции «связанных данных», поскольку каждому топониму не присвоен отдельный веб-адрес с постоянным и уникальным URL, что исключает возможность прямого гиперссылочного доступа к конкретным объектам.

### Обсуждение и заключение

В статье представлены такие топонимические сайты и базы данных, которые опубликованы в формате, который позволяет включать их в виде гиперссылок в Викиданные. Это существенно, поскольку Викиданные сейчас – эта самая большая и открытая база данных по всем отраслям наук. Таким образом, различные топонимические базы объединены в одну сеть связанных данных (Linked data). Это ведёт к популяризации этих топонимических баз данных, при этом данные становятся более доступными и удобными для научных исследований.

Объединение Викиданных и локальных баз данных позволяет расширить область поиска, улучшить качество данных за счёт их унифи-

кации и выявления возможных разночтений. Это также облегчает поддержку и разработку программ, работающих с такими базами, и способствует снижению количества ошибок. В результате создаётся более структурированная и надёжная система хранения и обработки информации.

Таким образом, данные проекта Топонимия Карелия доступны как связанные открытые данные. Публикуемые в дальнейшем наборы данных, включающие топонимы Карелии или Северо-Западного региона, могут применяться для объединения информации о культурном наследии Карелии и северо-запада России и приграничья путём описания музейных и архивных предметов с помощью гиперссылок из этих наборов данных на топонимы проекта ТопКар.

### Список источников и литературы

1. Захарова Е. В. ТопКар – геоинформационный ресурс по топонимии Карелии: историко-географический аспект // Исторический подход в географии и геоэкологии: Материалы VII Междун. науч.-образов. конф. по исторической географии. Петрозаводск: Изд-во ПетрГУ, 2023. С. 447–451.
2. Мартыненко И. А. Электронные ресурсы как инструменты для топонимических исследований // Вестник ТГПУ. 2020. № 5 (211). URL: <https://cyberleninka.ru/article/n/elektronnye-resursy-kak-instrumenty-dlya-toponimicheskikh-issledovaniy> (дата обращения: 01.04.2025).
3. Мищенко О. В. Электронная база данных «Топонимия Костромской области»: обработка и систематизация полевого материала // Вопросы ономастики. 2011. № 2 (11). С. 89–104.
4. Программирование Викиданных / А. А. Крижановский, М. С. Балакирева, Е. А. Меньшикова, Е. О. Паренченков, А. С. Потес, Е. Д. Трубина, А. Обрегон. Петрозаводск: Изд-во ПетрГУ, 2024. 186 с. URL: [https://commons.wikimedia.org/wiki/File:Wd\\_book\\_ru\\_2024.pdf](https://commons.wikimedia.org/wiki/File:Wd_book_ru_2024.pdf) (дата обращения: 01.04.2025).
5. Штерна М. База данных географических названий Латвии. UNGEGN, Вильнюс, 2000. URL: [https://web.archive.org/web/20240804111654/https://arhiiv.eki.ee/knn/ungegn/bd3lv\\_db.htm](https://web.archive.org/web/20240804111654/https://arhiiv.eki.ee/knn/ungegn/bd3lv_db.htm) (дата обращения: 10.09.2025).
6. Apenīte M., Bojārs U. National Library of Latvia subject headings as linked open data // The Semantic Web: ESWC 2021 Satellite Events. Cham: Springer, 2021. Pp. 33–37. DOI: 10.1007/978-3-030-80418-3\_6.
7. Drivers of participation in digital citizen science: Case studies on Järviwiki and Safecast / V. Palacin, S. Gilbert, S. Orchard, A. Eaton, M. A. Ferrario, A. Happonen // Citizen Science: Theory and Practice. 2020. Vol. 5. № 1. P. 22. DOI: 10.5334/cstp.290.
8. Entity Management Using Wikidata for Cultural Heritage Information / L. Zhu, A. Xu, S. Deng, G. Heng, X. Li // Cataloging & Classification Quarterly. 2023. Vol. 61. № 1. Pp. 20–46. URL: [https://projects.iq.harvard.edu/sites/projects.iq.harvard.edu/files/cbdb/files/entity\\_management\\_using\\_wikidata\\_for\\_cultural\\_heritage\\_information.pdf](https://projects.iq.harvard.edu/sites/projects.iq.harvard.edu/files/cbdb/files/entity_management_using_wikidata_for_cultural_heritage_information.pdf) (дата обращения: 01.04.2025).
9. GNIS-LD: Serving and Visualizing the Geographic Names Information System Gazetteer As Linked Data / B. Regalia, K. Janowicz, G. Mai, D. Varanka, E. L. Usery // The Semantic Web: 15<sup>th</sup> International Conference. ESWC 2018 (June 3–7, 2018, Heraklion). Crete, Greece. Pp. 528–540. URL: [https://gengchenmai.github.io/papers/2018-ESWC\\_GNIS-LD.pdf](https://gengchenmai.github.io/papers/2018-ESWC_GNIS-LD.pdf) (дата обращения: 01.04.2025).
10. Manning J., Turk A. How Terrain Becomes Landscape: Antarctica Landscape Language Case Study // Landscape Values Place and Praxis Conference. Ireland: Galway, 2016. Pp. 178–182. URL: [https://www.academia.edu/download/49841269/Landscape\\_Values\\_Place\\_and\\_Praxis\\_proceedings.pdf#page=179](https://www.academia.edu/download/49841269/Landscape_Values_Place_and_Praxis_proceedings.pdf#page=179) (дата обращения: 01.04.2025).
11. Mustonen T., Huusari N. How to know about waters? Finnish traditional knowledge related to waters and implications for management reforms // Reviews in Fish Biology and Fisheries. 2020. Vol. 30. № 4. Pp. 699–718. DOI: 10.1007/s11160-020-09619-7.
12. Registry of place names in Norwegian polar areas [Dataset]. Norwegian Polar Institute. 2014. DOI: 10.21334/NPOLAR.2011.A2813EB6.
13. Ormeling F. Evaluation of current international toponymic databases // Toponymy Training Manual. New York, 2017. URL: <https://unstats.un.org/unsd/geoinfo/ungegn/docs/CHAPTER%202021.pdf> (дата обращения: 10.09.2025).
14. Page R. Ten years and a million links: building a global taxonomic library connecting persistent identifiers for names, publications and people // Biodiversity Data Journal. 2023. Vol. 11. P. e107914. DOI: 10.3897/BDJ.11.e107914.

15. Sorrentino J. T. The New Monasticon Hibernicum and Inquiry into the Early Christian and Medieval Church in Ireland // Conference on Monasteries and Society in the Later Middle Ages, University of Wales, Aberystwyth and Lampeter, Gregynog Hall (Powys), 4–7 April 2005. URL: <https://www.york.ac.uk/media/borthwick/documents/publications/MRB10.pdf> (дата обращения: 01.04.2025).
16. Szczepańska J. Dolnosorbische mjenja | Niedersorbische Namen – internet service dedicated to Lower Sorbian proper names // *Onomastica*. 2022. Vol. 66. Pp. 329–340. DOI: 10.17651/ONOMAST.66.22.
17. Theo Van Veen. Wikidata: From “an” Identifier to “the” Identifier // *Information Technology and Libraries*. 2019. Vol. 38. Issue 2. Pp. 72–81. DOI: 10.6017/ital.v38i2.10886.
18. Vrandečić D., Pintscher L., Krötzsch M. Wikidata: The making of // *Companion Proceedings of the ACM Web Conference 2023*. 2023. Pp. 615–624. URL: <https://dl.acm.org/doi/pdf/10.1145/3543873.3585579> (дата обращения: 1.04.2025).
19. Zhang Z., Bethard S. A survey on geocoding: algorithms and datasets for toponym resolution // *Lang Resources & Evaluation*. 2024. DOI: 10.1007/s10579-024-09730-2.

## References

1. Zakharova E. V. *TopKar – geoinformatsionnyy resurs po toponimii Karelii: istoriko-geograficheskiy aspekt* [TopKar – geoinformation resource on the toponymy of Karelia: historical-geographical aspect]. *Istoricheskiy podkhod v geografii i geoekologii: Materialy VII Mezhdunarodnoy nauchno-obrazovatelnoy konferentsii po istoricheskoy geografii* [Historical Approach in Geography and Geoecology: Proceedings of the VII International Scientific and Educational Conference on Historical Geography]. Petrozavodsk: PetrSU Publ., 2023. Pp. 447–451. (In Russian)
2. Martynenko I. A. *Elektronnyye resursy kak instrumenty dlya toponimicheskikh issledovaniy* [Electronic resources as tools for toponymic research]. *Vestnik TGPU* [Tomsk State Pedagogical University Bulletin], 2020, no. 5 (211). Available at: <https://cyberleninka.ru/article/n/elektronnye-resursy-kak-instrumenty-dlya-toponimicheskikh-issledovaniy> (accessed April 01, 2025). (In Russian)
3. Mishchenko O. V. *Elektronnaya baza dannykh “Toponimiya Kostromskoy oblasti”: obrabotka i sistematizatsiya polevogo materiala* [The electronic database “Toponymy of the Kostroma Oblast”: processing and systematization of field data]. *Voprosy onomastiki* [Problems of Onomastics], 2011, no. 2 (11), pp. 89–104. (In Russian)
4. Krizhanovsky A. A., Balakireva M. S., Menshikova E. A., Parenchenkov E. O., Potes A. S., Trubina E. D., Obregón A. *Programmirovaniye Vikidannykh* [Programming Wikidata]. Petrozavodsk: PetrSU Publ., 2024. 186 p. Available at: [https://commons.wikimedia.org/wiki/File:Wd\\_book\\_ru\\_2024.pdf](https://commons.wikimedia.org/wiki/File:Wd_book_ru_2024.pdf) (accessed April 01, 2025). (In Russian)
5. Šterna M. *Baza dannykh geograficheskikh nazvaniy Latvii* [Latvian Geographical Names Database]. UNGEGN, Vilnius, 2000. Available at: [https://web.archive.org/web/20240804111654/https://arhiiv.eki.ee/knn/ungegn/bd3lv\\_db.htm](https://web.archive.org/web/20240804111654/https://arhiiv.eki.ee/knn/ungegn/bd3lv_db.htm) (accessed September 10, 2025). (In Russian)
6. Apenite M., Bojars U. *National Library of Latvia subject headings as linked open data*. The Semantic Web: ESWC 2021 Satellite Events. Cham: Springer, 2021. Pp. 33–37. DOI: 10.1007/978-3-030-80418-3\_6. (In English)
7. Palacin V., Gilbert S., Orchard S., Eaton A., Ferrario M. A., Happonen A. Drivers of participation in digital citizen science: Case studies on Järviwiki and Safecast. *Citizen Science: Theory and Practice*, 2020, no. 5 (1), p. 22. DOI: 10.5334/cstp.290. (In English)
8. Zhu L., Xu A., Deng S., Heng G., Li X. Entity Management Using Wikidata for Cultural Heritage Information. *Cataloging & Classification Quarterly*, 2023, no. 61 (1), pp. 20–46. Available at: [https://projects.iq.harvard.edu/sites/projects.iq.harvard.edu/files/cbdb/files/entity\\_management\\_using\\_wikidata\\_for\\_cultural\\_heritage\\_information.pdf](https://projects.iq.harvard.edu/sites/projects.iq.harvard.edu/files/cbdb/files/entity_management_using_wikidata_for_cultural_heritage_information.pdf) (accessed April 01, 2025). (In English)
9. Regalia B., Janowicz K., Mai G., Varanka D., Usery E. L. GNIS-LD: Serving and Visualizing the Geographic Names Information System Gazetteer As Linked Data. *Proceedings of ESWC 2018* (June 3–7, 2018, Heraklion). Crete, Greece, 2018. Pp. 528–540. Available at: [https://gengchenmai.github.io/papers/2018-ESWC\\_GNIS-LD.pdf](https://gengchenmai.github.io/papers/2018-ESWC_GNIS-LD.pdf) (accessed April 01, 2025). (In English)
10. Manning J., Turk A. How Terrain Becomes Landscape: Antarctica Landscape Language Case Study. *Landscape Values Place and Praxis Conference*. Ireland: Galway, 2016. Pp. 178–182. Available at: [https://www.academia.edu/download/49841269/Landscape\\_Values\\_Place\\_and\\_Praxis\\_proceedings.pdf#page=179](https://www.academia.edu/download/49841269/Landscape_Values_Place_and_Praxis_proceedings.pdf#page=179) (accessed April 01, 2025). (In English)
11. Mustonen T., Huusari N. How to know about waters? Finnish traditional knowledge related to waters and implications for management reforms. *Reviews in Fish Biology and Fisheries*, 2020, no. 30 (4), pp. 699–718. DOI: 10.1007/s11160-020-09619-7. (In English)
12. Registry of place names in Norwegian polar areas [Dataset]. *Norwegian Polar Institute*, 2014. DOI: 10.21334/NPOLAR.2011.A2813EB6. (In English)
13. Ormeling F. Evaluation of current international toponymic databases. *Toponymy Training Manual*. New York, 2017. Available at: <https://unstats.un.org/unsd/geoinfo/ungegn/docs/CHAPTER%2021.pdf> (accessed September 10, 2025). (In English)
14. Page R. Ten years and a million links: building a global taxonomic library connecting persistent identifiers for names, publications and people. *Biodiversity Data Journal*, 2023, no. 11, p. e107914. DOI: 10.3897/BDJ.11.e107914. (In English)

15. Sorrentino J. T. The New Monasticon Hibernicum and Inquiry into the Early Christian and Medieval Church in Ireland. *Conference on Monasteries and Society in the Later Middle Ages, University of Wales, Aberystwyth and Lampeter, Gregynog Hall (Powys) (4–7 April 2005)*. Available at: <https://www.york.ac.uk/media/borthwick/documents/publications/MRB10.pdf> (accessed April 01, 2025). (In English)
16. Szczepańska J. Dolnosorbische mjenja. Niedersorbische Namen – internet service dedicated to Lower Sorbian proper names. *Onomastica*, 2022, no. 66, pp. 329–340. DOI: 10.17651/ONOMAST.66.22. (In English)
17. Theo Van Veen. Wikidata: From “an” Identifier to “the” Identifier. *Information Technology and Libraries*, 2019, no. 38 (2), pp. 72–81. DOI: 10.6017/ital.v38i2.10886. (In English)
18. Vrandečić D., Pintscher L., Krötzsch M. Wikidata: The making of. *Companion Proceedings of the ACM Web Conference 2023*, 2023, pp. 615–624. Available at: <https://dl.acm.org/doi/pdf/10.1145/3543873.3585579> (accessed April 01, 2025). (In English)
19. Zhang Z., Bethard S. A survey on geocoding: algorithms and datasets for toponym resolution. *Lang Resources & Evaluation*, 2024. DOI: 10.1007/s10579-024-09730-2. (In English)

#### ИНФОРМАЦИЯ ОБ АВТОРЕ

**Крижановский Андрей Анатольевич**, ведущий научный сотрудник лаборатории информационных компьютерных технологий, Институт прикладных математических исследований КарНЦ РАН, ФИЦ «Карельский научный центр» (185910, Российская Федерация, Республика Карелия, г. Петрозаводск, ул. Пушкинская, д. 11), кандидат технических наук.

[andrew.krizhanovsky@gmail.com](mailto:andrew.krizhanovsky@gmail.com)  
ORCID ID: 0000-0003-3717-2079

#### ABOUT THE AUTHOR

**Krizhanovsky Andrey Anatolyevich**, Leading Researcher, Laboratory for Information Computer Technologies, Institute of Applied Mathematical Research, Karelian Research Center of the Russian Academy of Sciences (185910, Russian Federation, Republic of Karelia, Petrozavodsk, Pushkinskaya Str., 11), Candidate of Technical Sciences.

[andrew.krizhanovsky@gmail.com](mailto:andrew.krizhanovsky@gmail.com)  
ORCID ID: 0000-0003-3717-2079