Лабораторная работа 5 (уровень 1)

*Задания уровня 1*

1.1 «Этнология» (http://www.ethnologue.com/)

Представленная информация: содержит краткие сведения о 7102 живых языках мира с демографическими указаниями, функциональными характеристиками, картами распространения.

Для чего можно использовать: можно использовать этот сайт для изучения информации о различных языках мира, исследователи могут находить на этом сайте данные для своих исследований, студенты, преподаватели могут использовать этот ресурс для изучения разнообразия языков мира и их социокультурного значения.

Запросы: можно ввести запрос в поле ввода на главной странице, а также есть отдельная вкладка с языками и странами.

Преимущества: сайт предоставляет обширную информацию о более чем 7 000 языках мира, включая их классификацию, распространение, количество носителей, диалекты и другие характеристики, сайт регулярно обновляется, что позволяет получать актуальную информацию о языках и изменениях в их статусе.

Недостатки: платный доступ к полной информации: часть информации на сайте доступна только при подписке или платном доступе, что может быть ограничением для некоторых пользователей, возможные неточности: как и любая база данных, «Этнология» может содержать неточности или устаревшую информацию, поскольку языковая ситуация в мире постоянно меняется.

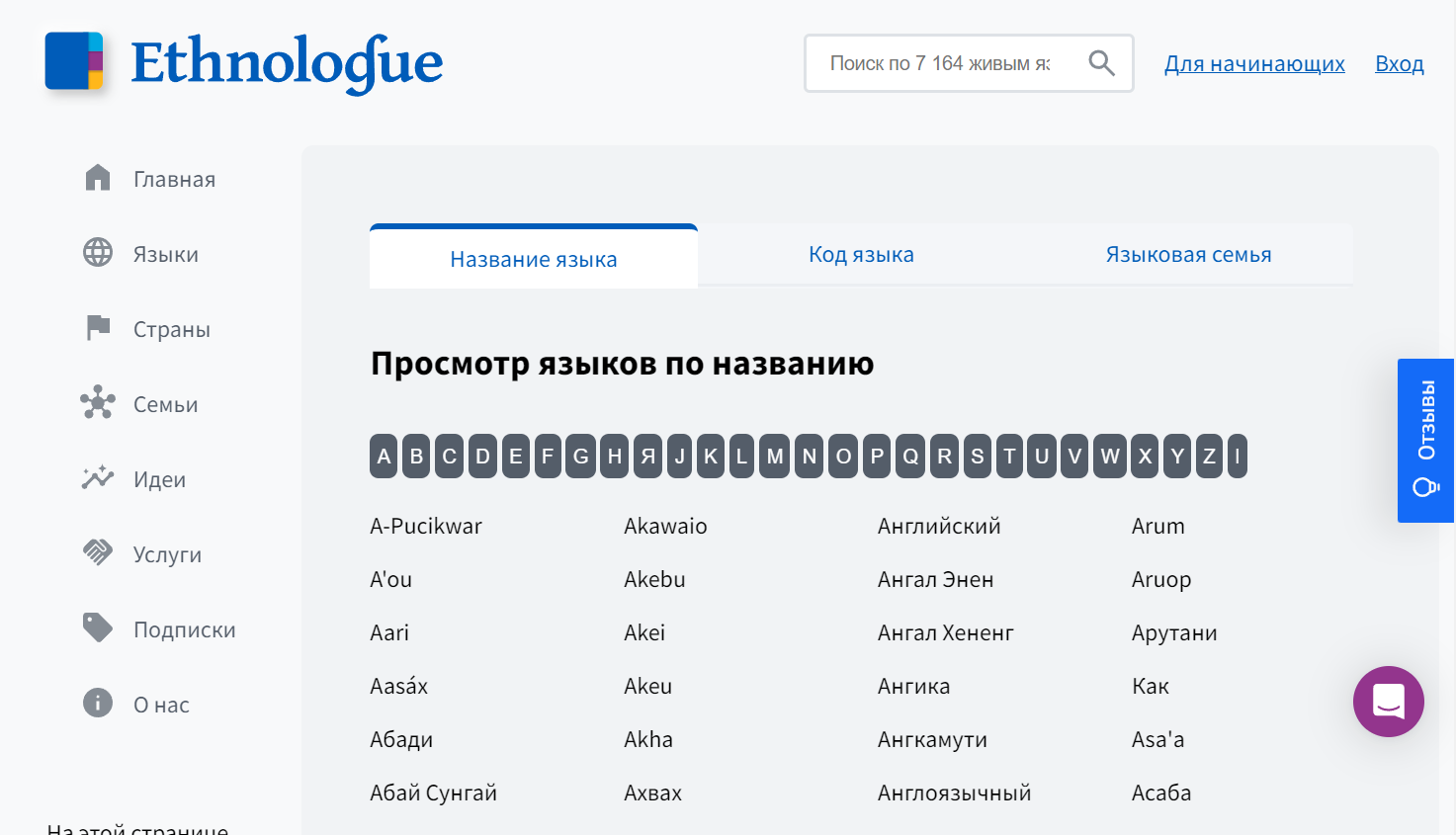
Похожие ресурсы:

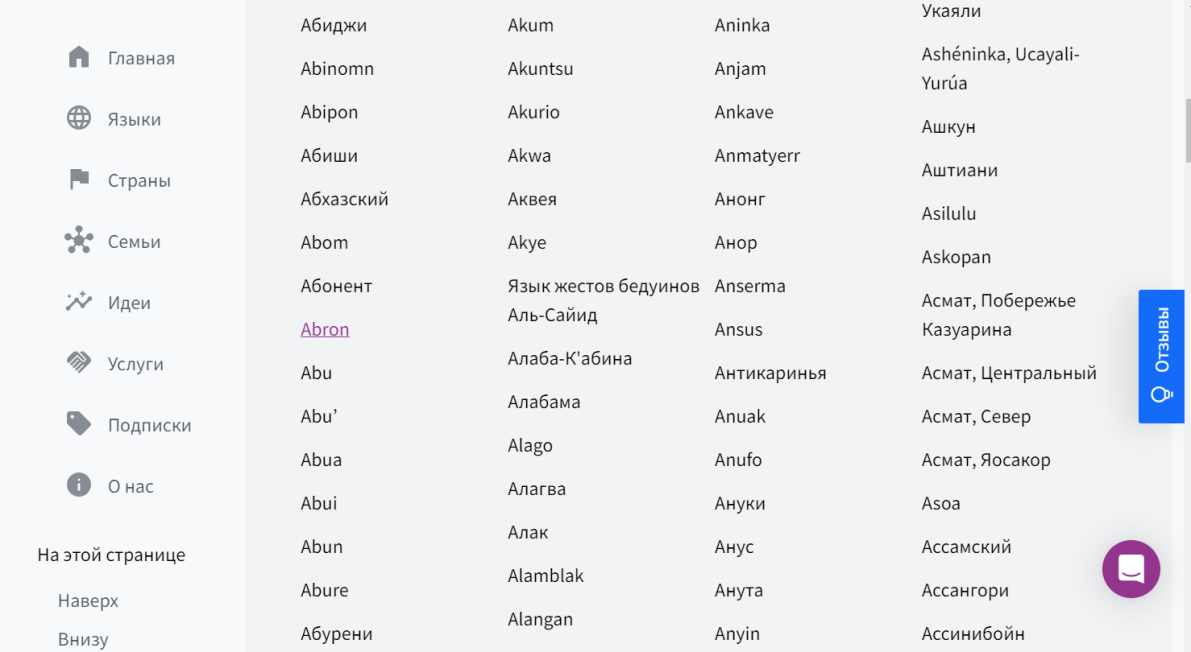
* Glottolog (https://glottolog.org/): Это база данных, которая содержит обширную информацию о языках, их классификации, родственных связях и лингвистических характеристиках.
* Linguasphere (http://www.linguasphere.info/): Этот ресурс предлагает классификацию языков мира и информацию о их распространении, сходстве и различиях.
* UNESCO Atlas of the World's Languages in Danger (http://www.unesco.org/languages-atlas/): Этот атлас предоставляет информацию о языках, находящихся под угрозой исчезновения, а также о мерах по сохранению и поддержке этих языков.

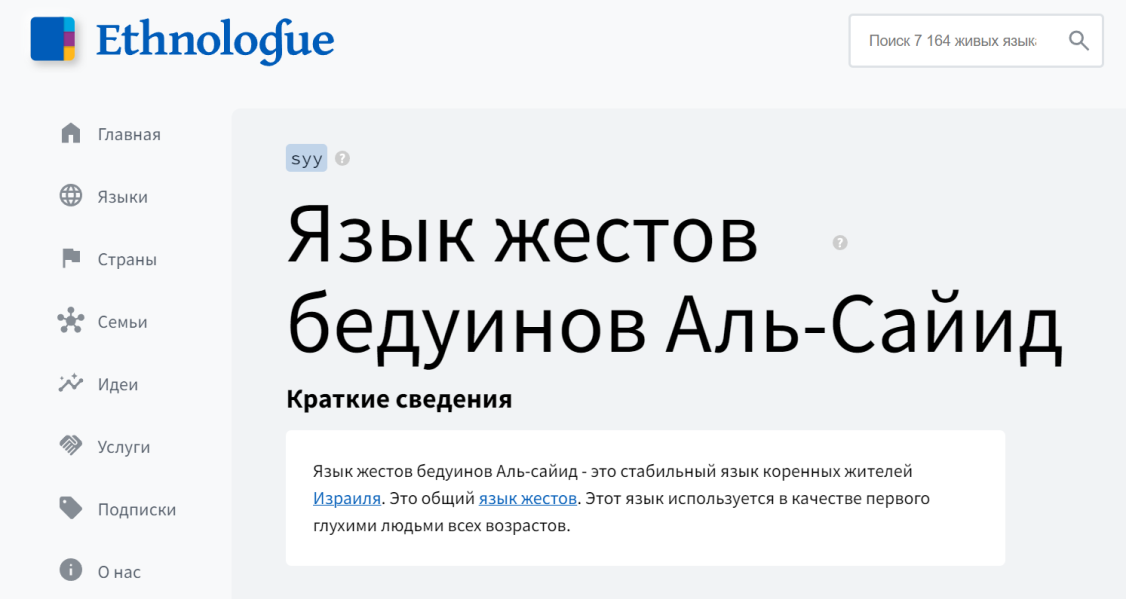
Исследования:

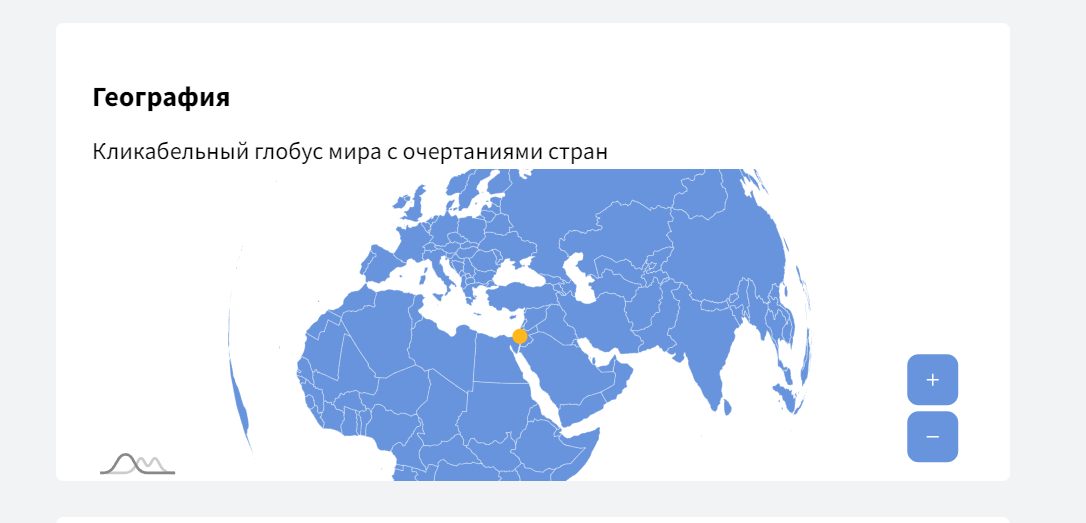
* Статистика языкового использования: исследование изменений в численности носителей языков, их распространения и статуса.
* Классификация языков: обновление и дополнение классификации языков по генетическим и лингвистическим критериям.
* Изучение угрожаемых языков: идентификация и документирование языков, находящихся под угрозой исчезновения.
* Исследование диалектов и вариантов языков: Анализ разнообразия диалектов и вариантов языков в различных регионах мира.

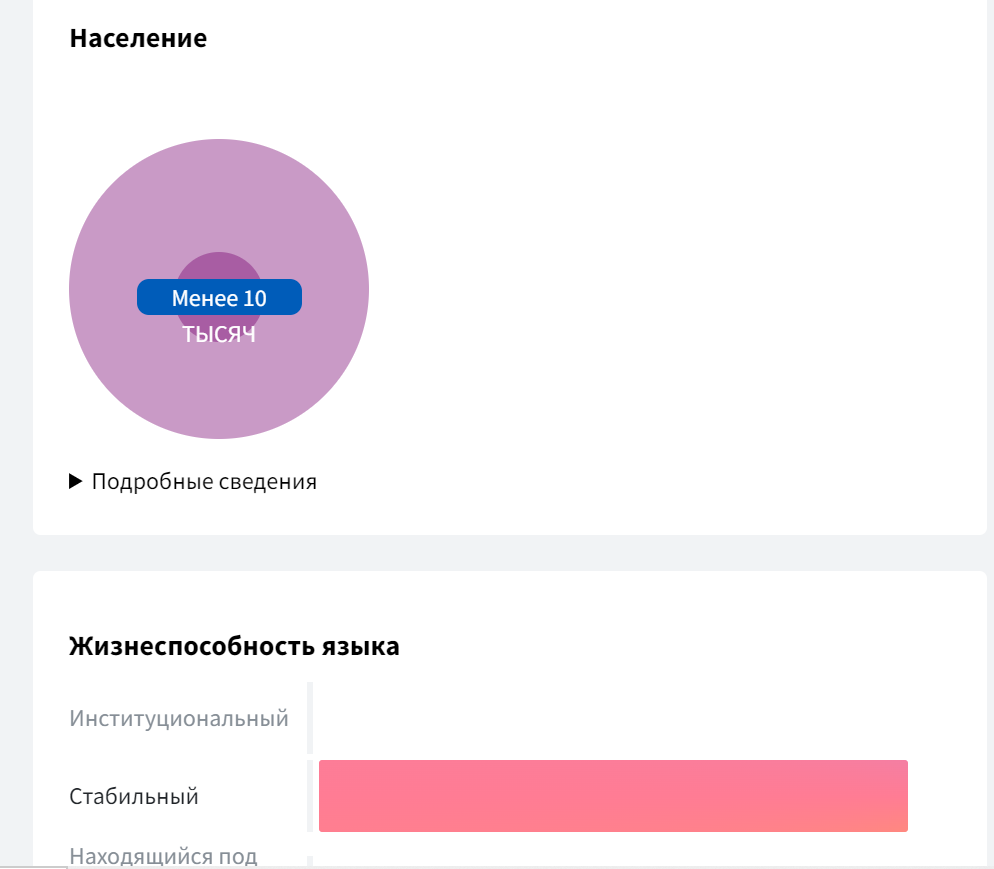
*Пример работы:*











Анализ: поиск выдал краткие сведения, население,расположение и жизнеспособность языка. Это позволяет узнать основную информацию про язык.

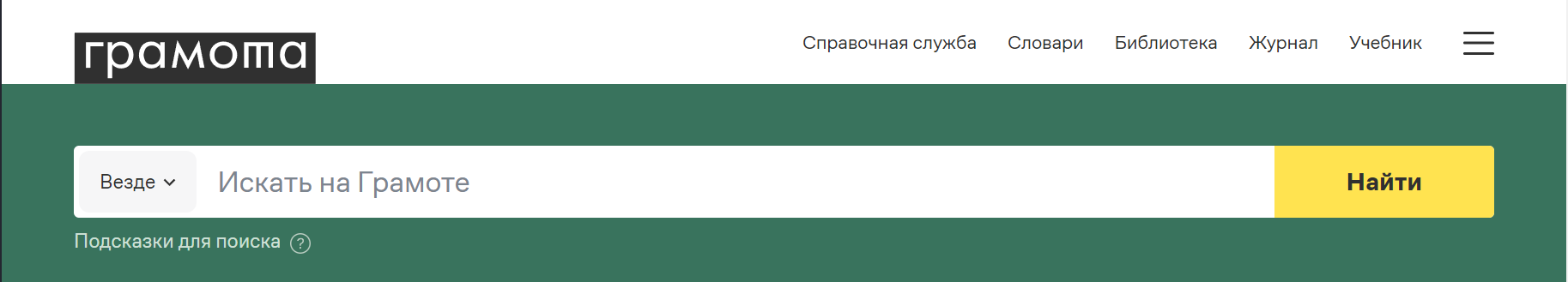
1.2 «Грамота.ру» (http://www.gramota.ru/)

Представленная информация: можно найти постоянно пополняющийся список ссылок на словари в Сети, в том числе на словари энциклопедические и терминологические.

Для чего можно использовать:

* Проверка правописания и грамматики: "Грамота.ру" предоставляет возможность проверить правописание слов, фраз, а также провести грамматический анализ текста.
* Словари и справочники: ресурс содержит различные словари, толковые словари, энциклопедии и другие справочники по русскому языку.
* Советы по стилистике и языку: здесь можно найти советы по стилистике, орфографии, пунктуации и другим аспектам русского языка.
* Обучающие материалы: "Грамота.ру" предлагает обучающие материалы для изучения русского языка, включая уроки, тесты и учебники.
* Форумы и консультации: на сайте можно задавать вопросы специалистам по русскому языку и получать консультации по различным языковым вопросам.

Запросы:



Преимущества:

* Проверка правописания и грамматики: "Грамота.ру" предоставляет возможность быстрой проверки правописания слов и фраз, а также анализа грамматических ошибок.
* Обширные словари и справочники: ресурс содержит различные словари, толковые словари и другие справочные материалы, что делает его полезным для расширения словарного запаса.
* Советы по стилистике и языку: "Грамота.ру" предлагает советы по стилистике, орфографии, пунктуации и другим аспектам русского языка, что помогает улучшить письменный стиль.
* Обучающие материалы: ресурс содержит обучающие материалы, уроки и тесты по русскому языку, что может быть полезно для изучающих язык или желающих улучшить свои знания.
* Форумы и консультации: на сайте есть возможность задавать вопросы экспертам по русскому языку и получать консультации, что помогает разрешить языковые вопросы.

Недостатки:

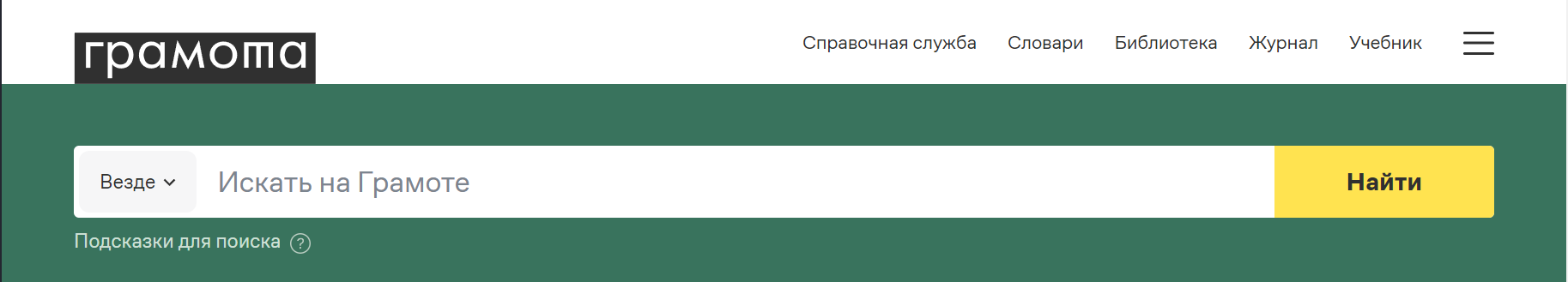
* Ограниченность информации: некоторые пользователи могут найти ограниченность информации на сайте по сравнению с более крупными ресурсами.
* Не всегда актуальная информация: информация на сайте может быть не всегда актуальной или не отражать последние изменения в правописании или грамматике.
* Ограниченные возможности: некоторые пользователи могут захотеть более широкий спектр функций, которые могут быть доступны на других ресурсах.

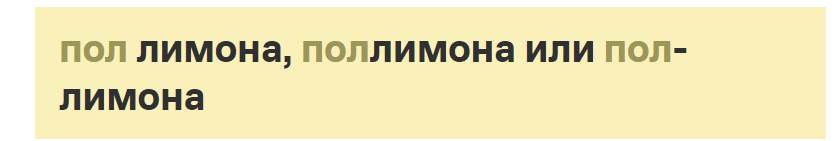
Похожие ресурсы:

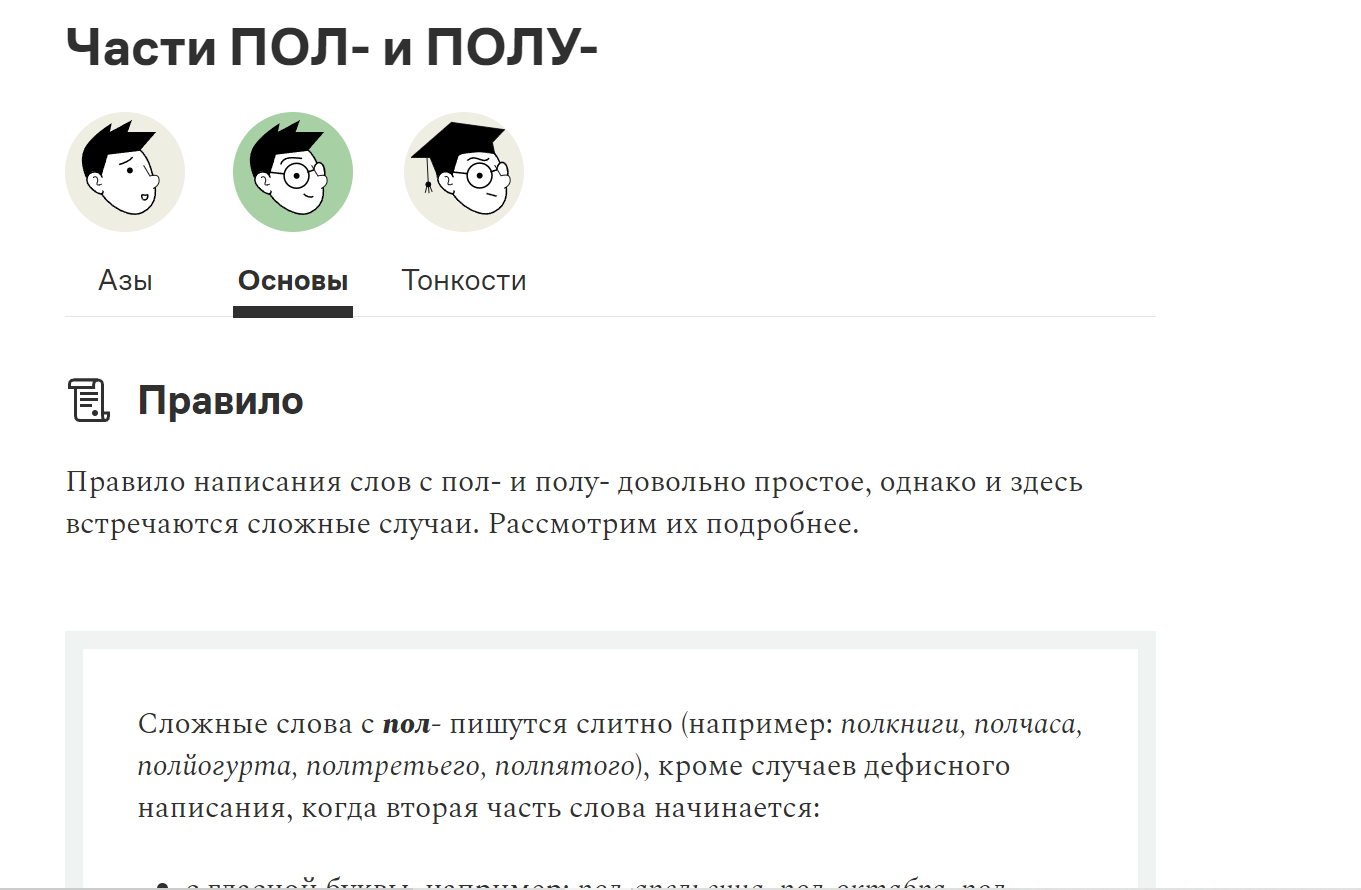
* "Русский Язык" (http://rus-yaz.ru/): этот ресурс предлагает различные материалы по русскому языку, включая правила грамматики, уроки, тесты и другие обучающие материалы.
* "Лингво.ру" (https://www.lingvo.ru/): Этот ресурс предоставляет онлайн-словари, переводчики, справочники по грамматике и другие инструменты для изучения и работы с русским языком.
* "Орфоэпия.ру" (https://orfogrammka.ru/): Этот сайт специализируется на орфоэпии — науке о произношении слов. Здесь можно найти правила произношения слов, упражнения и другие материалы.
* "Словари и энциклопедии на Академике" (https://www.academic.ru/): Этот ресурс предлагает доступ к различным словарям, энциклопедиям, толковым словарям и другим справочным материалам.

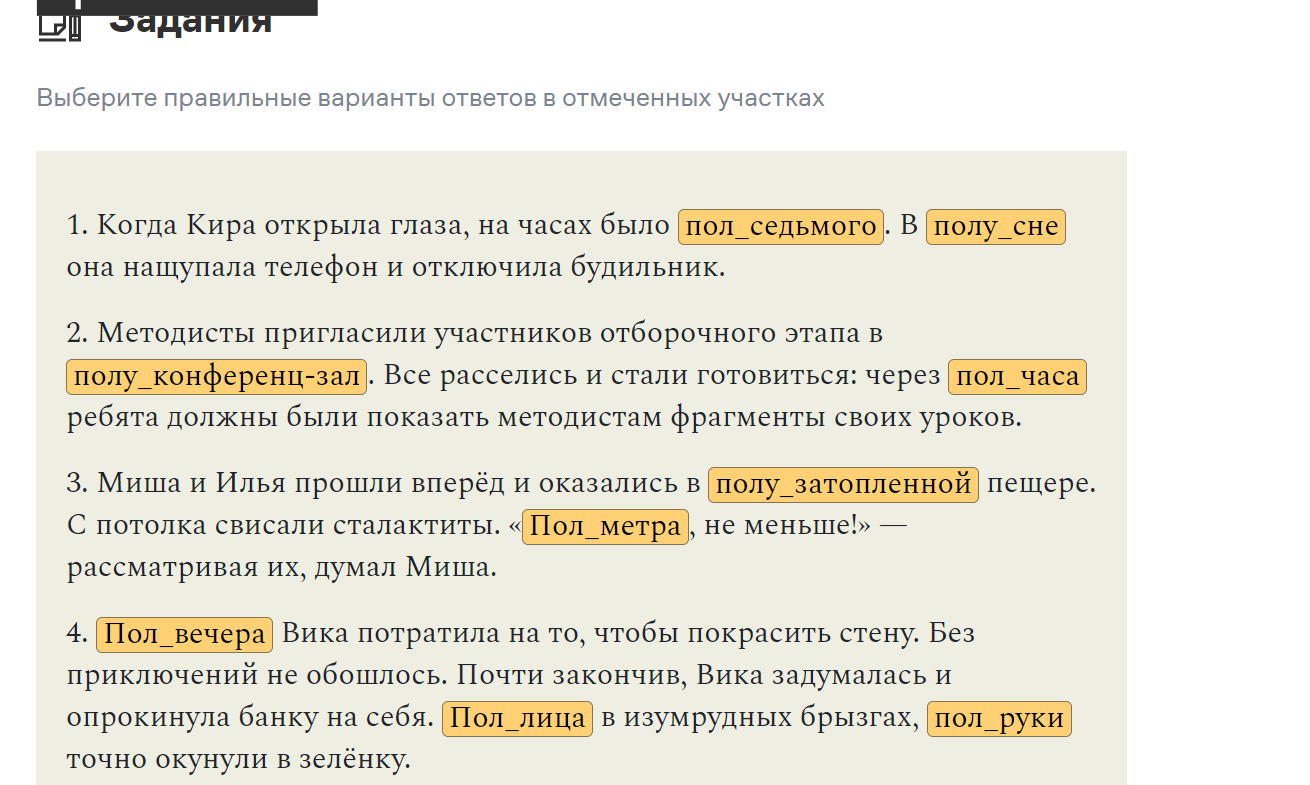
Исследования: на сайте не размещается информация о проведенных исследованиях.

*Пример работы:*









Анализ: поиск выдаёт полную информацию (правило, исключения, примечания), а также дает задания для практики.

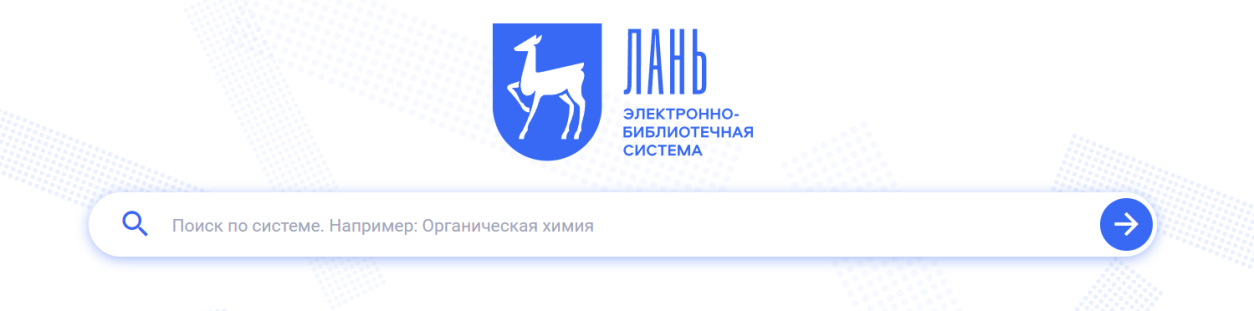
1.3  Электронно-библиотечная система издательства «Лань» (https://e.lanbook.com/)

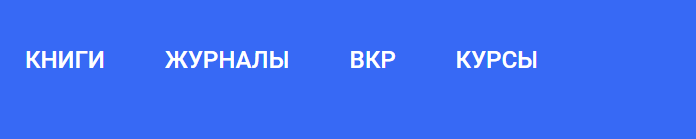
Представленная информация: ресурс, созданный с   целью   обеспечения   вузов   необходимой   учебной   и научной   литературой профильных   направлений.

Для чего можно использовать:

* Чтение книг: можно читать электронные книги прямо на сайте или скачивать их для чтения в оффлайн-режиме на устройствах, поддерживающих соответствующие форматы.
* Обучение и самообразование: электронно-библиотечная система может быть использована для обучения и самообразования. Вы можете найти книги по различным темам, которые помогут вам расширить свои знания.
* Исследования и научная работа: для студентов, ученых и исследователей доступ к электронным версиям книг может быть полезным для проведения исследований, написания научных работ и обзоров литературы.

Запросы:





Преимущества:

1. Широкий выбор книг: Система предоставляет доступ к большому количеству электронных книг издательства «Лань». Вы можете найти книги по различным темам и жанрам.

2. Удобство использования: Вы можете читать электронные книги прямо на сайте или скачивать их для чтения в оффлайн-режиме на устройствах, поддерживающих соответствующие форматы. Это делает использование системы удобным и гибким.

3. Экономия места: Электронные книги не занимают физического пространства, как печатные издания. Вы можете хранить большую коллекцию книг на одном устройстве.

4. Поиск и навигация: Система предоставляет возможность быстрого поиска и навигации по книгам. Вы можете легко найти нужную информацию, используя функции поиска, закладки и гиперссылки.

5. Удобство для путешествий: Если вы часто путешествуете, электронная библиотека позволяет брать с собой множество книг без необходимости носить их физически.

Недостатки:

1. Ограниченный выбор: Возможно, система не предлагает такой широкий выбор книг, как другие электронные библиотеки или печатные магазины.

2. Зависимость от технологии: Для использования системы требуется наличие устройства (компьютера, планшета или смартфона) и доступа в Интернет. Если у вас нет подходящего устройства или отсутствует интернет-соединение, вы не сможете получить доступ к книгам.

3. Ограничения в использовании: Возможно, система имеет ограничения на количество одновременно доступных книг или срок пользования определенными книгами.

4. Отсутствие физического взаимодействия: Некоторым людям нравится физическое взаимодействие с книгами, перелистывание страниц и запах бумаги. Этого аспекта может не хватать при чтении электронных книг.

5. Охрана авторских прав: Некоторые электронные библиотеки имеют ограничения на копирование, печать и распространение содержимого книг из соображений защиты авторских прав.

Похожие ресурсы:

1. LitRes (https://www.litres.ru/): LitRes — это крупнейшая электронная библиотека и магазин электронных книг на русском языке. Здесь можно найти широкий выбор книг различных жанров.

2. Bookmate (https://bookmate.com/): Bookmate предлагает доступ к огромной коллекции электронных книг на разных языках. Пользователи могут читать книги онлайн или скачивать их для оффлайн-чтения.

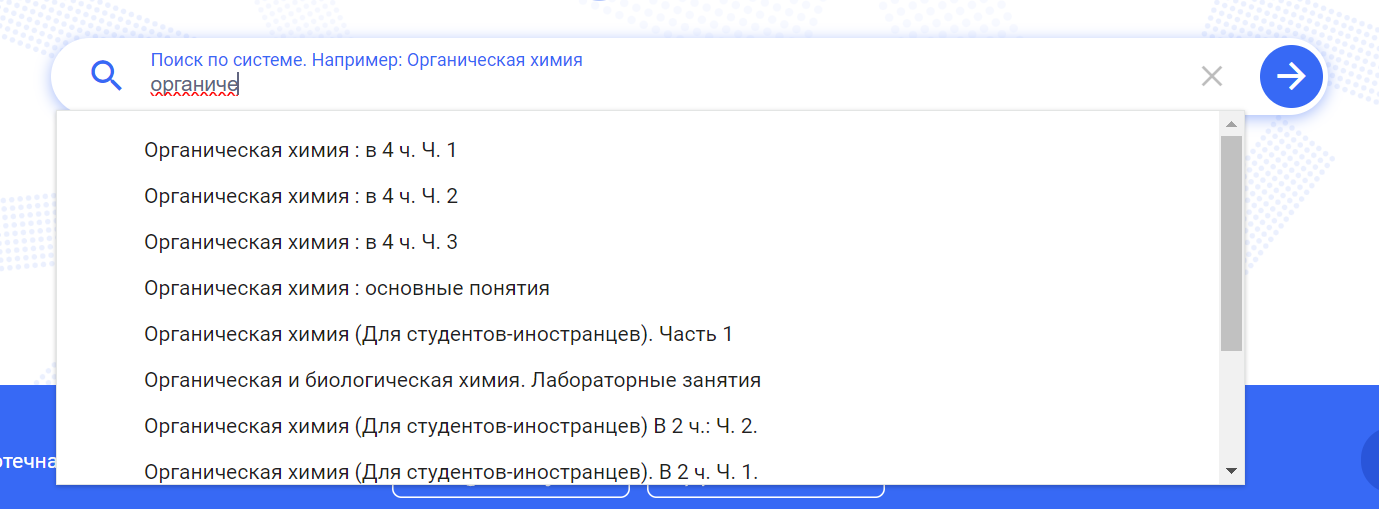
3. LitMarket (https://litmarket.ru/): Этот сервис предоставляет доступ к электронным книгам различных издательств. Здесь можно найти как бестселлеры, так и книги молодых авторов.

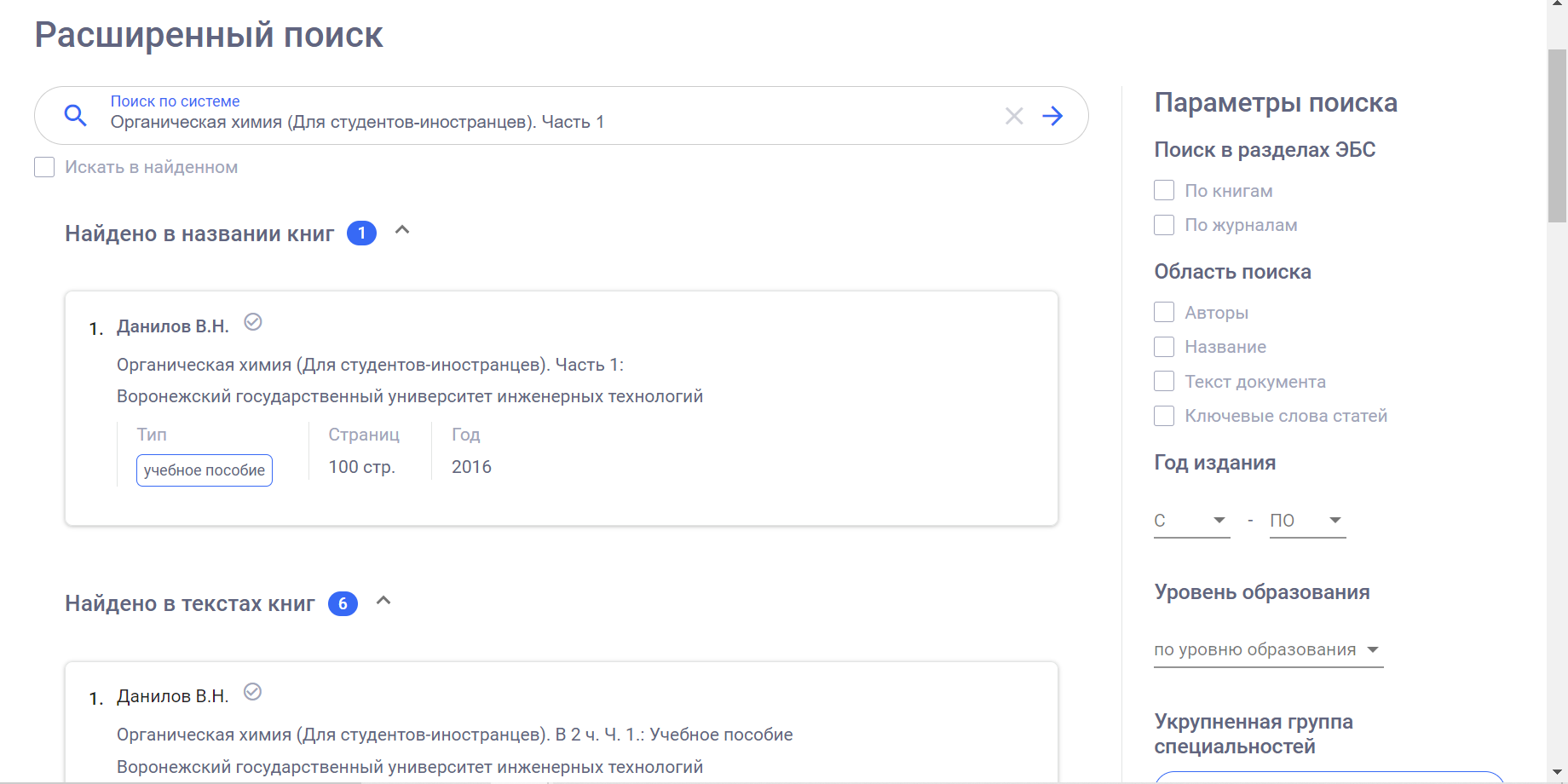
4. Amazon Kindle Store (https://www.amazon.com/kindle-books): Amazon Kindle Store предлагает огромный выбор электронных книг на английском и других языках. Вы можете приобрести книги для чтения на устройствах Kindle или через приложение Kindle на других устройствах.

5. Google Play Books (https://play.google.com/store/books): Google Play Books — это еще один популярный ресурс для приобретения и чтения электронных книг. Здесь можно найти книги различных жанров и авторов.

Исследования: не удалось найти

*Пример работы:*





Анализ: есть расширенный поиск, дана полная информация о книге, можно задать параметры.

1.4 Информационный блог об изучении иностранных языков (http://wt-blog.net/ )

Представленная информация: данный портал для тех,

кто хочет изучать иностранные языки и интересуется профессиями

переводчика и филолога. В блоге обычно можно найти полезные статьи, советы, рекомендации, методики и другую информацию, связанную с изучением иностранных языков.

Для чего можно использовать:

* Обучение иностранным языкам: блог может предоставлять полезные советы, методики и ресурсы для эффективного изучения иностранных языков. Это может включать в себя рекомендации по выбору курсов, учебников, онлайн-ресурсов и приложений.
* Получение мотивации: чтение успехов других людей в изучении иностранных языков, а также статей о преодолении трудностей, может помочь вдохновить и мотивировать вас на изучение языка.
* Получение информации о культуре: изучение иностранного языка часто включает изучение культуры страны, где этот язык является официальным. Блог может содержать материалы о культуре, традициях, праздниках и т.д.
* Обмен опытом: читатели могут делиться своим опытом изучения языков, обсуждать проблемы, находить решения и помогать друг другу.

Запросы: можно перейти по ссылке **«Все статьи»**, если нужны **все материалы в хронологическом порядке** или можно изучить **разделы портала**ниже на странице. Чтобы найти конкретную статью или материал, нужно воспользоваться **«Поиском по сайту»**или поиском в боковой колонке сайта.

Преимущества:

* Полезные советы и ресурсы: блог предоставляет полезные советы, методики и ресурсы для эффективного изучения иностранных языков, что может быть очень ценно для начинающих и опытных изучающих.
* Разнообразие тем: блог охватывает широкий спектр тем, связанных с изучением языков, что позволяет читателям получить разностороннюю информацию.
* Вдохновляющий контент: блог может содержать истории успеха, мотивационные материалы и другой контент, который помогает мотивировать читателей на изучение иностранных языков.
* Возможность общения: наличие комментариев или форума на сайте позволяет читателям общаться между собой, делиться опытом и задавать вопросы.

Недостатки:

* Устаревшая информация: если блог не обновляется регулярно, то информация на нем может устареть, что снижает его ценность для читателей.

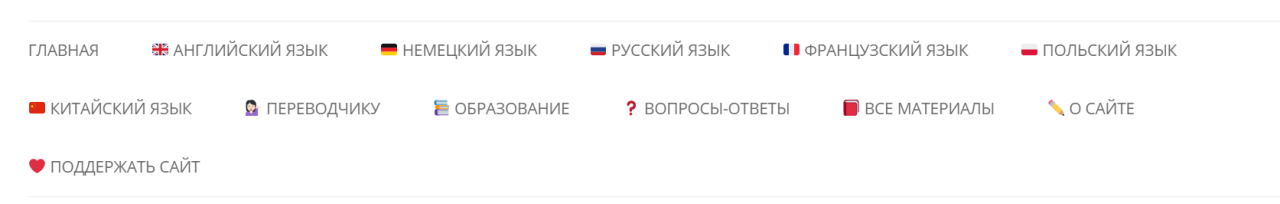
Похожие ресурсы:

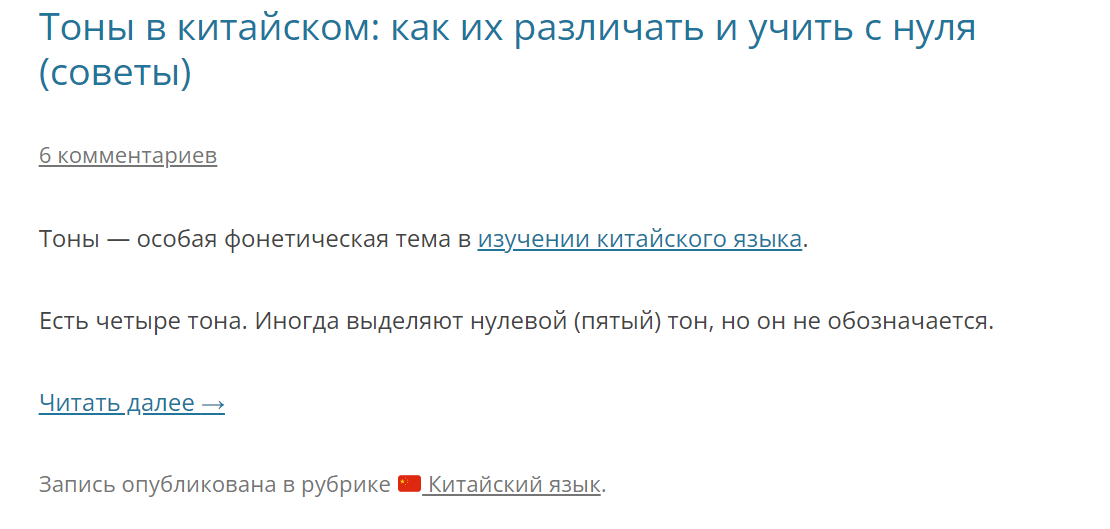
* LingQ: блог и сообщество для изучающих языки, где пользователи могут читать тексты, смотреть видео и слушать аудиозаписи на изучаемом языке, а также общаться с другими изучающими.
* Fluent in 3 Months: блог и подкаст, в которых Бенни Льюис делится своими советами и методами по быстрому изучению иностранных языков.
* Language Hacking: блог и YouTube-канал, где Скотт Янг рассказывает о своих экспериментах по изучению языков и делится практическими советами по эффективному изучению.
* Polyglot Club: онлайн-сообщество для полиглотов и изучающих языки, где можно найти советы, ресурсы и поддержку.
* Duolingo: популярное мобильное приложение для изучения иностранных языков с помощью игровых элементов.
* Babbel: платформа с фокусом на разговорной речи и повседневной коммуникации на иностранных языках.
* FluentU: сервис, который использует видео- и аудиоматериалы для изучения языков.
* italki: онлайн-платформа для обучения иностранным языкам с помощью преподавателей из разных стран.

Исследования: общие темы и направления исследований, которые часто проводятся в области изучения иностранных языков:

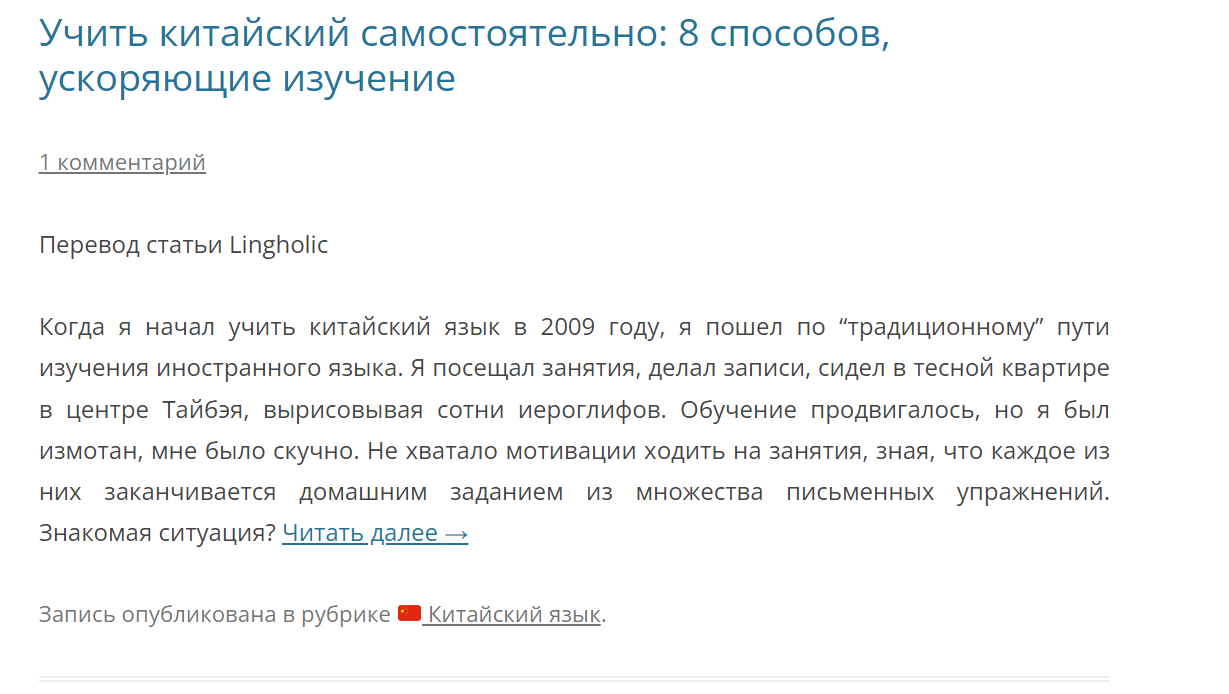
* Эффективные методики изучения иностранных языков.
* Влияние технологий на процесс обучения иностранным языкам.
* Психологические и когнитивные аспекты изучения иностранных языков.

*Пример работы:*









Анализ: поиск выдает полезные материалы и статьи, что помогает в изучении языка.

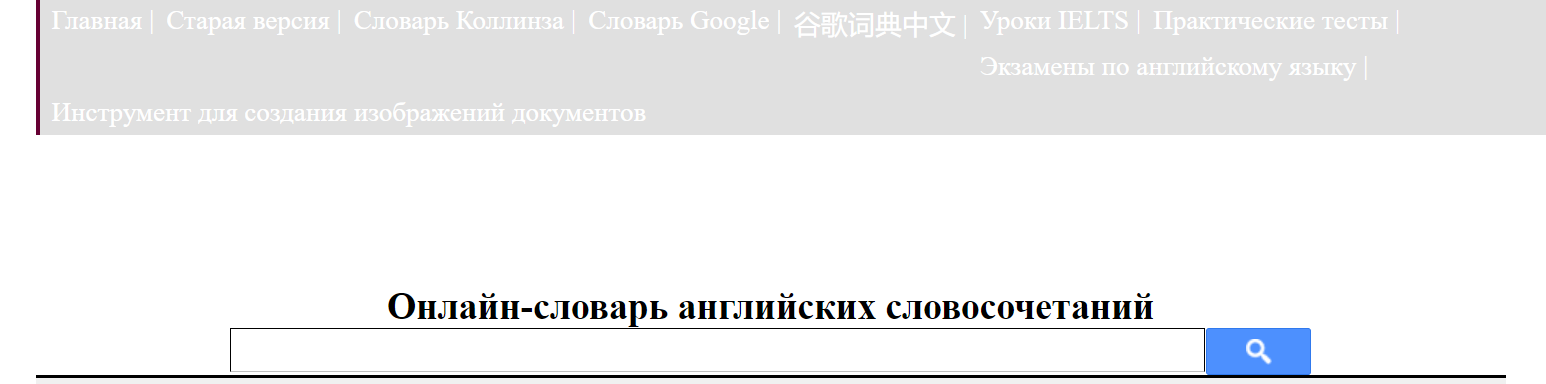
1.5  Онлайн - словарь  Oxford  Collocation  (http://www.freecollocation.com/)

Представленная информация: словари коллокаций предоставляют информацию о типичных сочетаниях слов в языке. Они помогают изучающим язык лучше понять, какие слова часто используются вместе и как правильно использовать их в контексте.

Для чего можно использовать:

* Изучение коллокаций: можно использовать словарь, чтобы изучать типичные сочетания слов в английском языке. Это поможет грамотно и естественно строить фразы и предложения.
* Расширение словарного запаса: онлайн-словарь Oxford Collocation может помочь вам расширить свой словарный запас, предоставляя различные словосочетания, которые могут быть использованы в разговорной и письменной речи.
* Помощь в написании: при написании эссе, статей или других текстов вам может потребоваться найти наиболее подходящие коллокации для передачи определенного значения или оттенка.
* Изучение английского языка: для тех, кто изучает английский как второй язык, использование онлайн-словаря Oxford Collocation может помочь улучшить понимание того, какие слова часто используются вместе, что способствует лучшему овладению языком.

Запросы:



Преимущества:

* Бесплатный доступ: ресурс предоставляет бесплатный доступ к информации о коллокациях, что делает его доступным для широкого круга пользователей.
* Удобство использования: веб-сайт прост в использовании и предлагает удобный интерфейс для поиска коллокаций по ключевым словам.
* Полезные сведения: Oxford Collocation предоставляет ценные сведения о типичных словосочетаниях, что помогает пользователям грамотно строить предложения и тексты на английском языке.

Недостатки:

* Ограниченный объем информации: некоторые пользователи могут обнаружить, что ресурс не содержит всей необходимой информации о коллокациях или ограничен в своем объеме.
* Не всегда актуальная информация: информация о коллокациях может быть не всегда актуальной или полной, так как язык постоянно развивается, и могут появляться новые сочетания слов.
* Отсутствие контекста: при использовании онлайн-словаря не всегда предоставляется контекст, в котором используется данное словосочетание, что может затруднить понимание его правильного использования.

Похожие ресурсы:

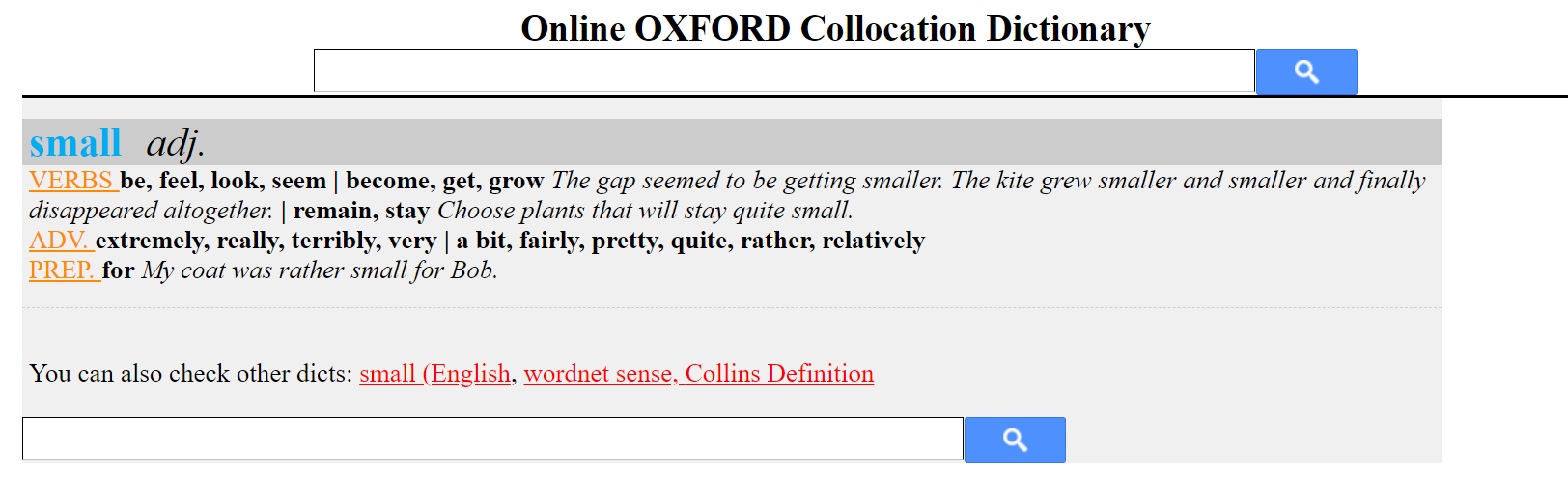
* Cambridge English Dictionary (https://dictionary.cambridge.org/): онлайн-словарь Cambridge также предоставляет информацию о коллокациях и типичных словосочетаниях на английском языке. Он является авторитетным источником для изучения английского языка.
* Macmillan Collocations Dictionary (https://www.macmillandictionary.com/collocations/): Macmillan Collocations Dictionary специализируется на предоставлении информации о типичных словосочетаниях и коллокациях в английском языке.
* WordReference (https://www.wordreference.com/): WordReference — это популярный онлайн-словарь, который также может помочь в изучении коллокаций и различных сочетаний слов в английском языке.
* Thesaurus.com (https://www.thesaurus.com/): Thesaurus.com предлагает синонимы и антонимы, что может быть полезно при изучении различных словосочетаний и их вариаций.

Исследования:

* Изучение влияния коллокаций на понимание текста: исследуется, как использование определенных коллокаций влияет на понимание текста читателями.
* Сравнительный анализ коллокаций в разных языках: исследование сравнивает типичные словосочетания и коллокации в разных языках для выявления сходств и различий.

*Пример работы:*

Поле ввода: small



Анализ: поиск предоставил информацию о типичных сочетаниях слова в языке.

*2. Вопросы для обсуждения:*

* Что может являться единицей корпуса?

Единицей крпуса могут являться тексты или их достаточно значительные фрагменты.

* Как отбираются тексты для корпуса? Проиллюстрируйте принципы отбора на примере Брауновского и других корпусов.

Представительность: тексты должны представлять весь спектр языка, включая различные регистры, жанры, периоды времени и т.д.

Качество: тексты должны быть хорошо написаны, без грамматических или орфографических ошибок.

Доступность: тексты должны быть доступны для сбора и обработки.

Объем: корпус должен быть достаточно большим, чтобы обеспечить статистическую значимость для проведения анализа и обучения моделей.

1. Брауновский корпус:

- Представительность: брауновский корпус был создан с целью представить различные стили и жанры английского языка. Он включает тексты из различных источников, таких как газеты, журналы, романы, научные статьи и т.д., чтобы обеспечить широкий охват языковых особенностей.

- Разнообразие: корпус содержит тексты различных жанров, от новостей до художественной литературы, чтобы обеспечить разнообразие языковых стилей и контекстов.

- Объем: брауновский корпус включает более миллиона слов, что обеспечивает достаточный объем для лингвистических исследований и обучения моделей.

2. Penn Treebank:

- Представительность: Penn Treebank включает аннотированные тексты для лингвистических исследований, такие как фразовая структура предложений, части речи и другие лингвистические аспекты.

- Разнообразие: корпус содержит тексты из газет, журналов, транскрибированных разговоров и других источников для обеспечения разнообразия контекстов.

- Объем: Penn Treebank включает большое количество аннотированных предложений для обучения и тестирования языковых моделей.

3. Google Books Ngram Corpus:

- Представительность: этот корпус включает данные из миллиардов книг, чтобы отразить употребление слов и выражений в течение длительного временного периода.

- Разнообразие: Google Books Ngram Corpus охватывает широкий спектр литературных жанров и тематик, от классики до современной литературы.

- Объем: корпус содержит огромное количество текстовых данных, что обеспечивает статистическую значимость для анализа употребления слов и выражений.

Каждый из этих корпусов был создан с учетом определенных целей и принципов отбора текстов, чтобы обеспечить представительность, разнообразие, объем и качество данных для соответствующих лингвистических исследований и задач обработки естественного языка.

* Представьте классификацию корпусов. Поясните, что означает «исследовательский корпус», «статический корпус», «параллельный корпус».

1. Исследовательский корпус:

- Описание: исследовательский корпус создается специально для конкретного исследования или задачи в области лингвистики или обработки естественного языка.

- Цель: используется для проведения экспериментов, анализа языковых особенностей, проверки гипотез и других научных целей.

2. Статический корпус:

- Описание: статический корпус представляет собой набор текстов, который сохраняется в неизменном виде и не обновляется.

- Цель: используется для долгосрочных исследований, анализа изменений в языке со временем или для обучения и тестирования моделей на стабильном наборе данных.

3. Параллельный корпус:

- Описание: параллельный корпус содержит тексты на двух или более языках, которые являются переводами друг друга.

- Цель: используется для машинного перевода, анализа синтаксических соответствий между языками, изучения межъязыковых отношений и других задач, связанных с переводом и сопоставлением языков.

*Корпусы языка могут быть классифицированы по различным критериям,*

*таким как:*

1. Объем данных:

- Маленькие корпусы: содержат относительно небольшое количество

текстовых данных, обычно используются для специфических исследований.

- Большие корпусы: включают в себя огромное количество текстов, что

позволяет проводить более обширные исследования и анализы.

2. Тип текстов:

- Общие корпусы: содержат разнообразные тексты из различных

источников и жанров.

- Специализированные корпусы: содержат тексты из определенной области

знаний или жанра (например, медицинские тексты, юридические документы

и т.д.).

3. Источник данных:

- Собственные корпусы: созданные исследователями или организациями

для конкретных целей.

- Открытые корпусы: доступные для общего пользования и собранные из

различных источников, таких как интернет, газеты, книги и т.д.

4. Языковая пара:

- Монолингвальные корпусы: содержат тексты только на одном языке.

- Параллельные корпусы: содержат тексты на двух или более языках,

соответствующие друг другу.

5. Структура данных:

- Текстовые корпусы: содержат только текстовую информацию.

- Мультимодальные корпусы: включают в себя различные типы данных,

такие как текст, звук, видео и т.д.

6. Цель использования:

- Обучающие корпусы: используются для обучения и разработки моделей

машинного обучения.

- Исследовательские корпусы: созданы для лингвистических исследований

и анализа языка.

* Выберите один из корпусов из списка ниже и охарактеризуйте его по следующим критериям: количество словоупотреблений, вид корпуса (по разным признакам).

Британский национальный корпус (British National Corpus - BNC):

Количество словоупотреблений: содержит около 100 миллионов словоупотреблений.

Вид корпуса:

- Собранный корпус: собран специально для исследования английского языка.

- Текстовый корпус: содержит только текстовую информацию.

- Общий корпус: охватывает широкий спектр жанров и тематик.

* Составьте глоссарий по теме «Корпусная лингвистика». Используйте для этого рекомендуемые источники литературы и сетевые ресурсы. Включите в глоссарий определения следующих понятий: конкорданс, рандомизация, коллокация, подмассив, парсинг, лемматизация, корпус-менеджер.

1. Конкорданс:

- Определение: Список или таблица, показывающая контексты, в которых встречается определенное слово или выражение в тексте корпуса, обычно с указанием окружающих слов.

- Источник: [Oxford Reference](https://www.oxfordreference.com/view/10.1093/oi/authority.20110803095438676)

2. Рандомизация:

- Определение: Процесс случайного выбора элементов из корпуса для создания случайной выборки или для исключения влияния порядка текстов на анализ.

- Источник: [ResearchGate](https://www.researchgate.net/publication/220887603\_Randomization\_in\_Corpus\_Linguistics)

3. Коллокация:

- Определение: Статистическая связь между словами, которая показывает, что они часто встречаются вместе в тексте и образуют стабильные выражения.

- Источник: [Cambridge Dictionary](https://dictionary.cambridge.org/dictionary/english/collocation)

4. Подмассив:

- Определение: Набор данных, содержащий подмножество текстов из корпуса, обычно используемый для конкретного анализа или исследования.

- Источник: [ResearchGate](https://www.researchgate.net/publication/229049004\_Corpus\_Linguistics\_and\_Statistical\_Analysis)

5. Парсинг:

- Определение: Анализ предложения на естественном языке с целью определения его структуры и выделения значимых элементов.

- Источник: [Stanford NLP Group](https://nlp.stanford.edu/IR-book/html/htmledition/stemming-and-lemmatization-1.html)

6. Лемматизация:

- Определение: Процесс приведения словоформы к её базовой форме (лемме) для упрощения анализа текста и поиска.

- Источник: [Natural Language Toolkit (NLTK)](https://www.nltk.org/book/ch05.html)

7. Корпус-менеджер:

- Определение: Программное обеспечение, предназначенное для управления и анализа корпусов текстов, включая поиск, фильтрацию, анализ и визуализацию данных.

- Источник: [LancsBox Software](https://corpora.lancs.ac.uk/lancsbox/)

* Найдите сетевые ресурсы по теме «корпусная лингвистика» и кратко охарактеризуйте их.

1. Sketch Engine (https://www.sketchengine.eu/):

- Описание: Sketch Engine представляет собой веб-платформу для работы с корпусами текстов на различных языках. Она предлагает широкий спектр инструментов для анализа корпусов, включая поиск словосочетаний, создание конкордансов, извлечение коллокаций и многое другое.

2. Corpus of Contemporary American English (COCA) (https://www.english-corpora.org/coca/):

- Описание: COCA является одним из самых крупных и наиболее авторитетных корпусов английского языка. Он содержит миллионы текстовых примеров, собранных из различных источников, таких как газеты, журналы, телевизионные программы и т.д. COCA позволяет исследователям анализировать лексические, грамматические и стилистические особенности современного американского английского.

3. Natural Language Toolkit (NLTK) (https://www.nltk.org/):

- Описание: NLTK представляет собой платформу и набор инструментов для обработки естественного языка на языке Python. Он содержит модули для работы с корпусами текстов, включая функции для токенизации, лемматизации, построения частотных распределений и других задач корпусной лингвистики.

2.1

|  |  |  |  |
| --- | --- | --- | --- |
|  | НКРЯ | COCA | BNC |
| Количество вхождений | 66-151 | 200 | 169 |

Вывод: больше всего вхождений у COCA, меньше всего у НКРЯ.

2.2

|  |  |  |  |
| --- | --- | --- | --- |
| № примера | НКРЯ | COCA | BNC |
| 1 | **Наталия Черникова. От школы до академии // «Наука и жизнь», 2009**  Когда появились в *русском* языке слова — школа, гимназия, лицей, институт, университет, академиям откуда они к нам пришли? | SPOK: ABC News: This Week  Not only that, but we know that the Republican party is afraid that without *Russian*interference, this President can't win. | One step backwards, two steps forward. Pethybridge, Roger. Oxford: Clarendon Press, 1990, pp. 22-144. 1797 s-units.  For the most part Great Russians were fighting Great Russians, and in the process trampled over the lands of peasants, whether Russian, Belorussian, Ukrainian, etc., none of whom was more inclined thereafter to feel particularly attracted to Muscovite patriotism. |
| 2 | **Максим Горький. Старуха Изергиль (1895)**  — Не хочу, — ответил я ей. — У!.. стариками родитесь вы, *русские*. Мрачные все, как демоны… | ACAD: Foreign Affairs  It was subject to sudden *Russian* whims. | The Daily Mirror. 9610 s-units.  # AN Aeroflot jet crashed in fog in the **Russian** city of Ivanovo, killing all 82 people on board, it was revealed yesterday. |
| 3 | **С. Соколова. Вечные истории // «Кот Шрёдингера», 2016**  Ну и *русская* классика, конечно, Дубровский успел бы оградить Машу от замужества; Хлестаков не выдал бы себя за ревизора; майору Ковалёву пришили бы нос — был бы краше прежнего. | MAG: Ars Technica  The two-reactor, 70MW floating power plant is headed through the Baltic Sea and north around Norway, to a *Russian*town called Murmansk, where the boat will receive its fuel. | The Daily Mirror. 7680 s-units.  But Pallister found the nerve-racking task too much in the pouring rain of the **Russian** capital. |

2.3

|  |  |  |  |
| --- | --- | --- | --- |
| № примера | НКРЯ | COCA | BNC |
| 1 | прил | прил | прил |
| 2 | прил | прил | прил |
| 3 | прил | прил | прил |

2.4

|  |  |  |  |
| --- | --- | --- | --- |
| № примера | НКРЯ | COCA | BNC |
| 1 | Язык | Страна | Люди |
| 2 | Люди | Люди | Относящийся к России |
| 3 | Относящаяся к русским | Относящийся к России | Относящийся к России |

2.5

1. Морфологические характеристики остаются постоянными для данного слова в разных корпусах.

2. Лексические характеристики (такие как значение, семантика, употребление в определенных контекстах) могут различаться в разных корпусах, так как они зависят от конкретного контекста, в котором слово употребляется.

2.6

1. Лексикография: корпусы могут использоваться для составления словарей, тезаурусов и других лексикографических ресурсов.

2. Семантические исследования: корпусы позволяют изучать значения слов, их употребление в различных контекстах.

3. Исследования стилистики и стилометрии: корпусы могут использоваться для анализа стилей авторов, выявления стилистических особенностей текстов, а также для автоматической стилометрии (определения авторства текстов).

4. Исследования языковых изменений: корпусы позволяют отслеживать изменения в языке со временем, изучать эволюцию лексики, грамматики и семантики.

5. Машинный перевод и обработка естественного языка: корпусы используются для обучения моделей машинного перевода.