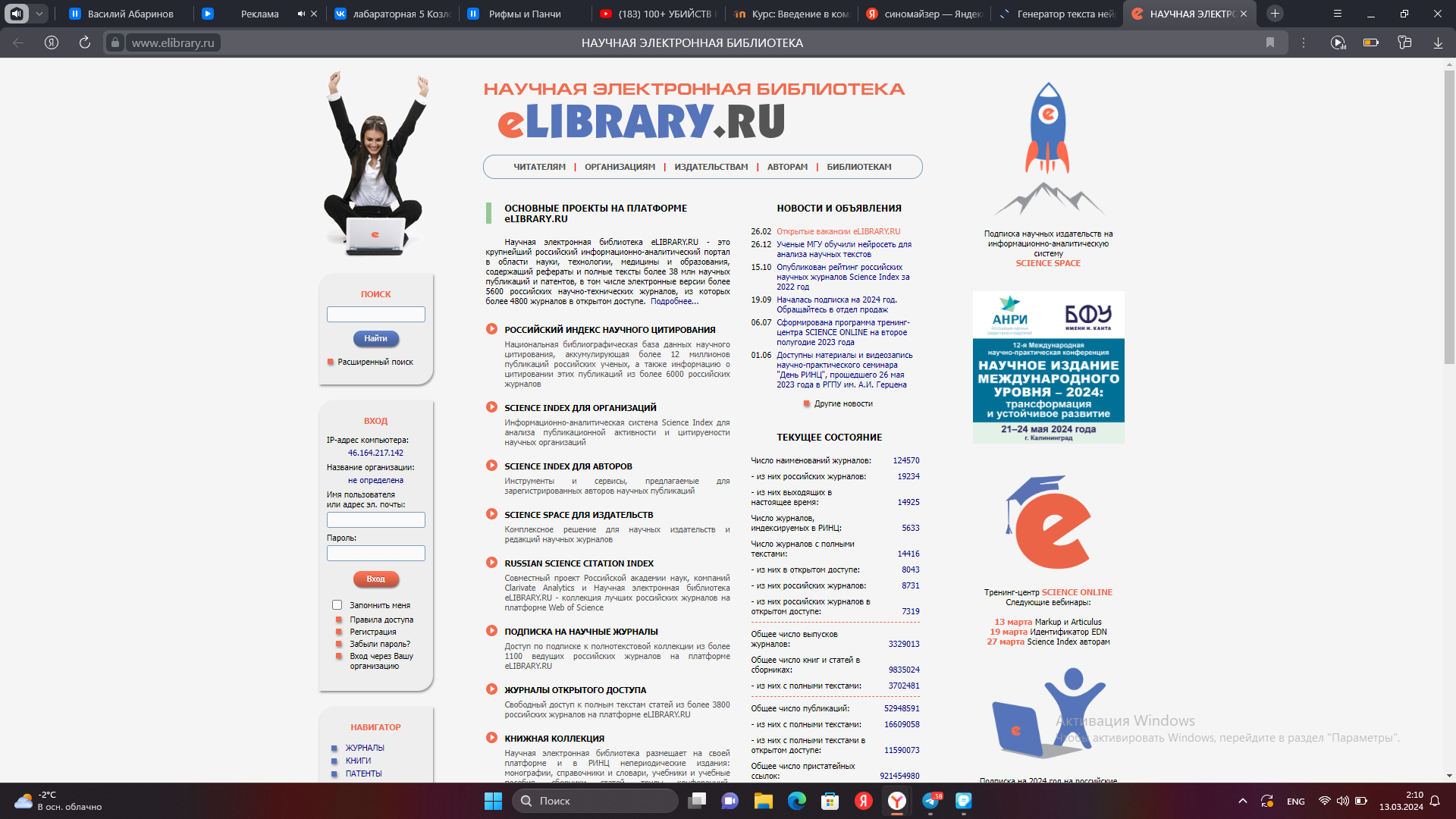
1. **ELIBRARY.RU (<https://elibrary.ru/defaultx.asp>?)**

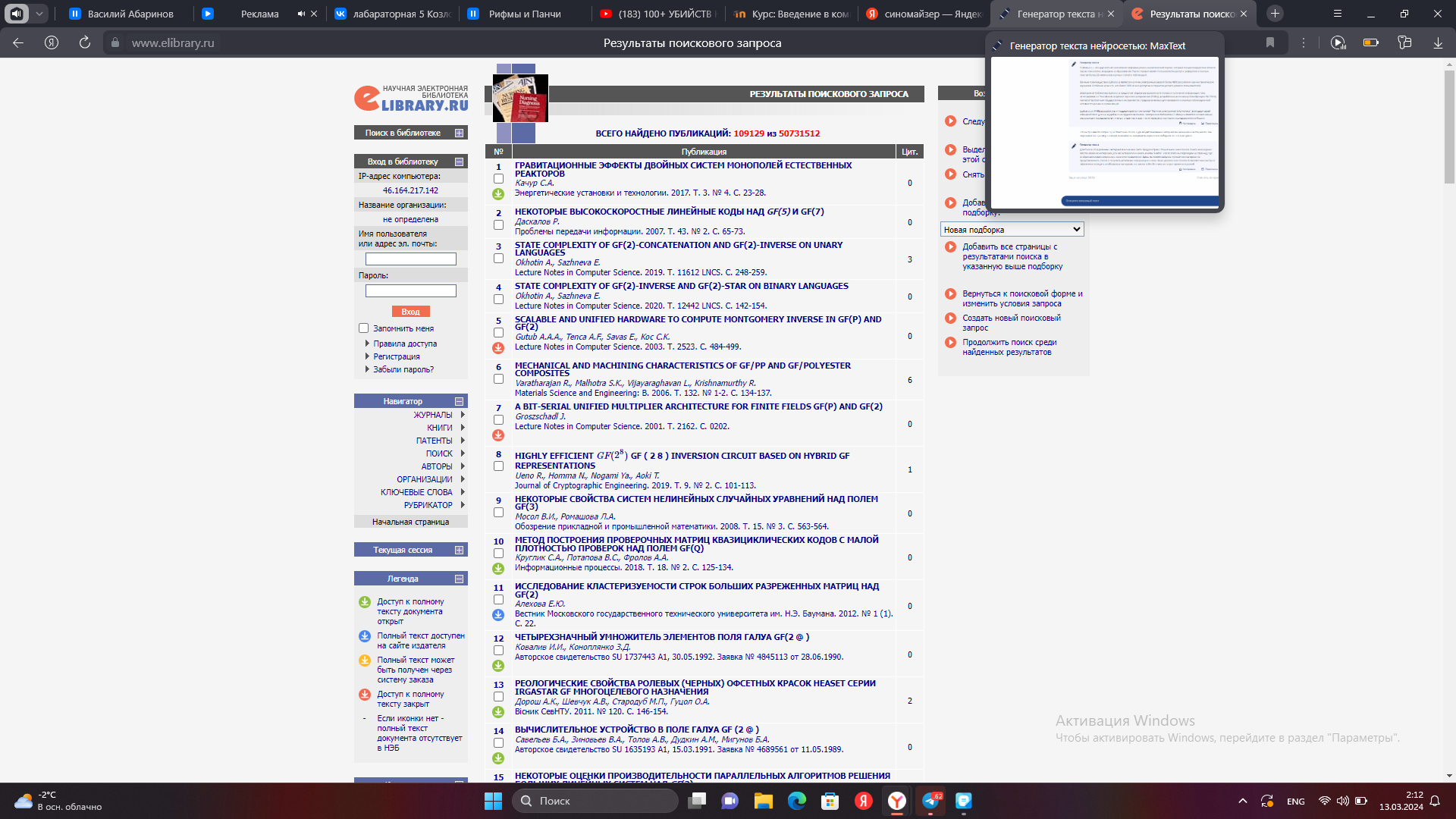
ELibrary.ru — это крупнейший российский информационно-аналитический портал, который специализируется в области науки, технологии, медицины и образования. Портал предоставляет пользователям доступ к рефератам и полным текстам более 22 миллионов научных статей и публикаций.Важным преимуществом ELibrary.ru является наличие электронных версий более 4800 российских научно-технических журналов. Особенно ценно то, что более 3800 из них доступны в открытом доступе для всех пользователей.Электронная библиотека ELibrary.ru предлагает обширные возможности поиска и получения информации. Она интегрирована с Российским индексом научного цитирования (РИНЦ), разработанным по заказу Минобрнауки РФ. РИНЦ является бесплатным общедоступным инструментом, предназначенным для измерения и анализа публикационной активности ученых и организаций.

ELibrary.ru и РИНЦ разработаны и поддерживаются компанией "Научная электронная библиотека". Благодаря своей обширной базе данных и удобным инструментам поиска, электронная библиотека ELibrary.ru является незаменимым ресурсом для исследователей, ученых, студентов и всех, кто интересуется научными исследованиями в России.

"The effects of mindfulness meditation on stress reduction in college students" - автор: Brown, S.

"The relationship between physical activity and mental health in elderly adults" - автор: Thompson, M.

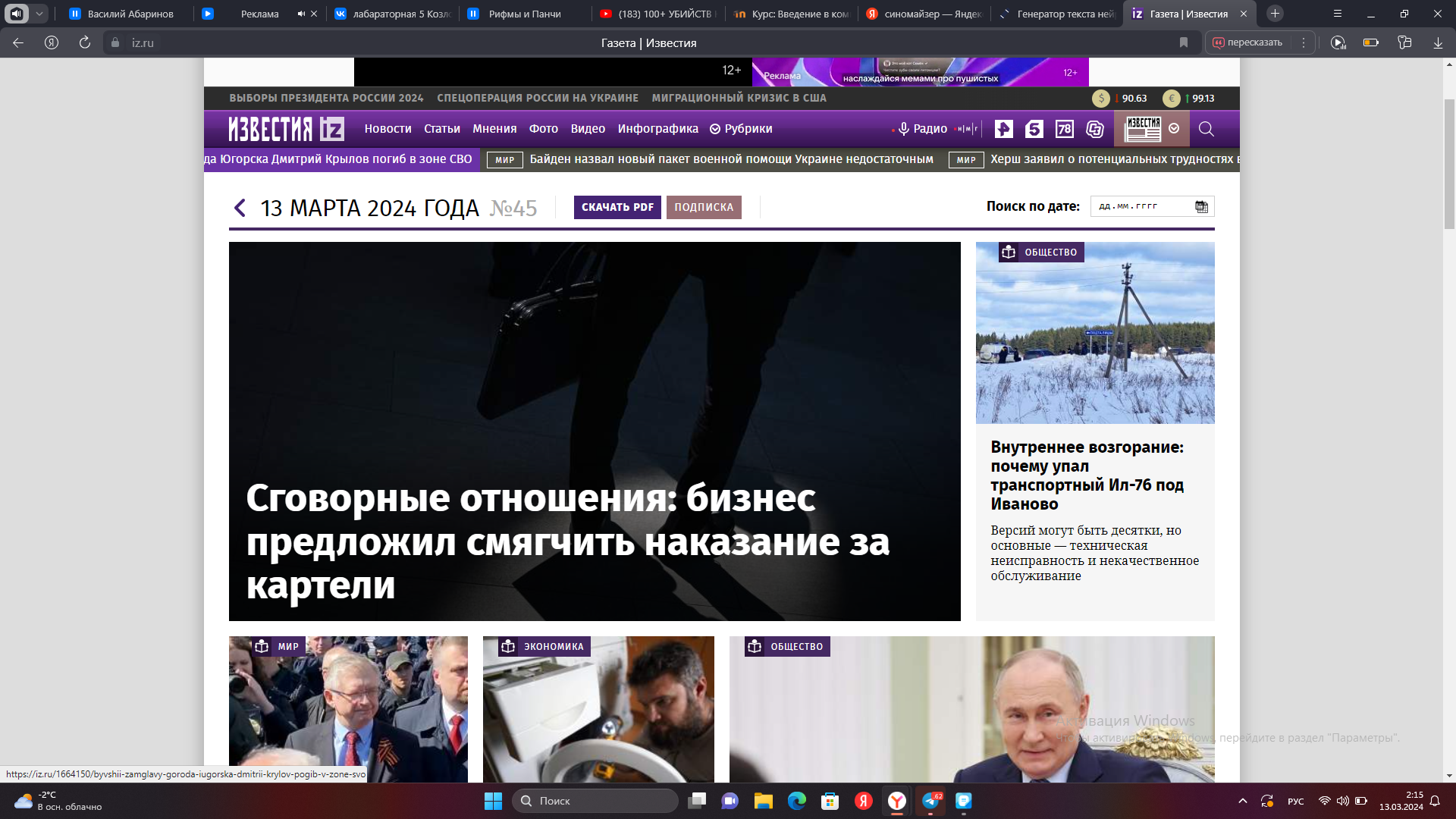
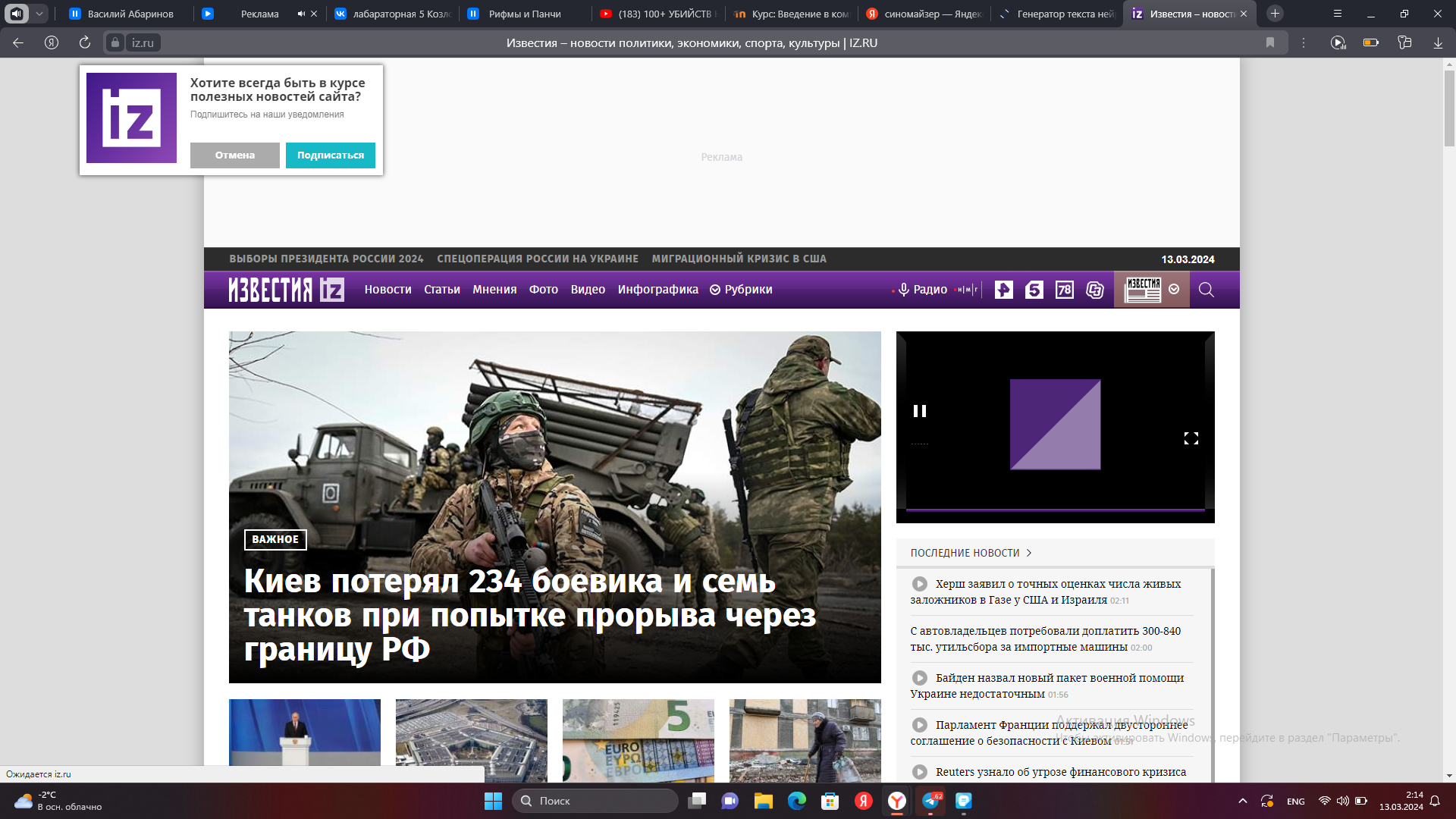


Для поиска определенных материалов на нашем сайте предусмотрено специальное окно поиска. В него необходимо ввести название интересующего нас материала и нажать кнопку "Найти". После этого мы переходим на страницу, где отображаются все материалы с похожими названиями. Здесь мы можем выбрать нужный нам материал из представленного списка и получить детальную информацию о нем. Наше удобное окно поиска позволяет нам быстро и эффективно находить необходимые материалы на нашем сайте без лишних затрат времени и усилий.

1. Известия (<https://iz.ru/>)

Лента новостей, обновляемая в режиме реального времени публикациями о событиях политики, экономики, социальной жизни, культуры, спорта и др.

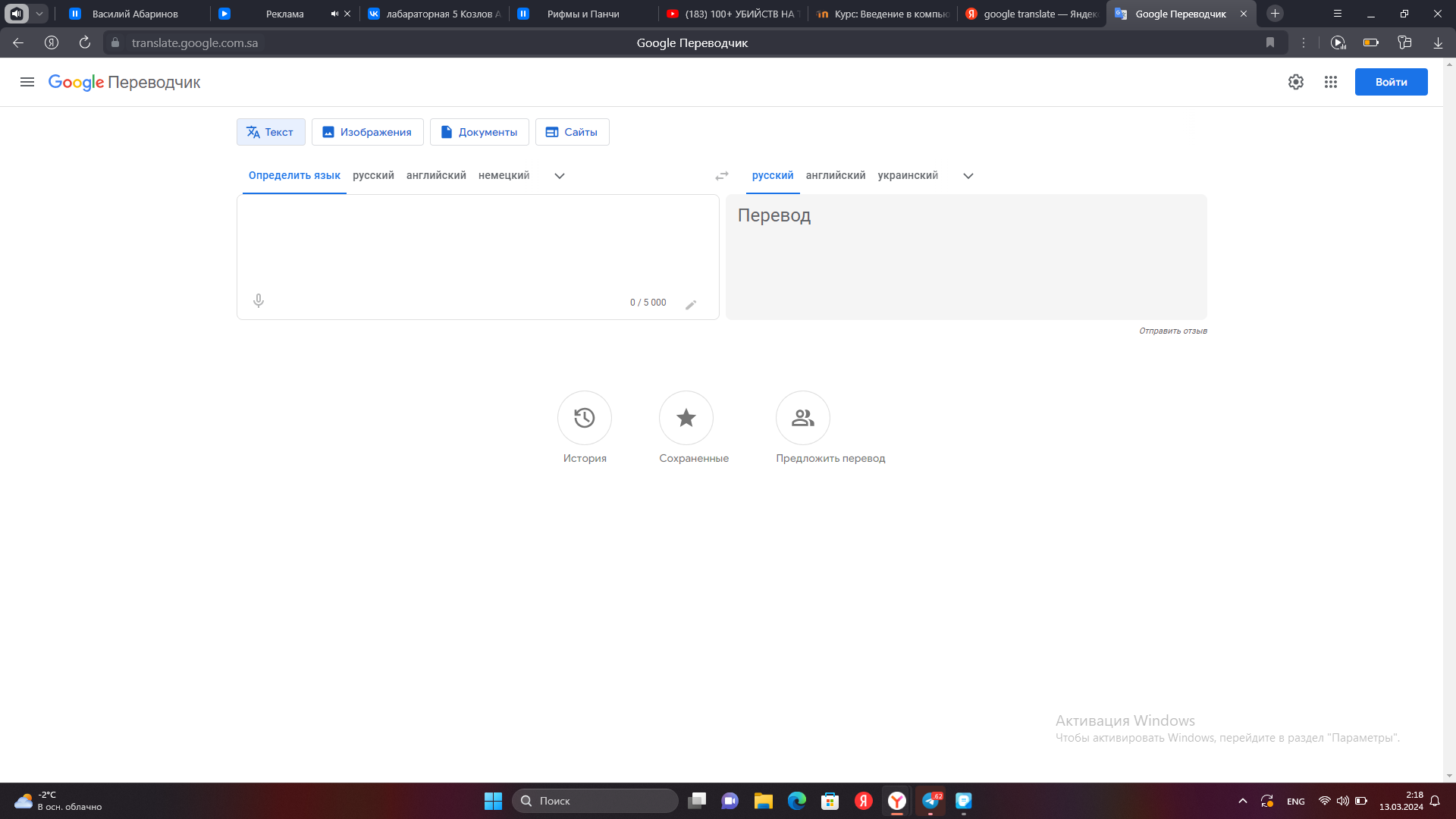
Этот ресурс можно использовать для получения необходимой информации о каких-либо событиях. В нем можно найти информацию о том, что происходит в мире, о достижениях спортсменов и тд. В нем содержаться фото, видео, статьи, новости, мнения экспертов и не только о чем-либо, всевозможные рубрики. Чтобы осуществить запрос нужно щелкнуть по значку лупы в правом верхнем углу и сделать запрос о той или иной теме, после чего вы попадаете на страницу, где вам представлены всевозможные статьи об этой теме



Преимущества этого ресурса состоят в том, что он очень прост в эксплуатации, можно всегда найти свежую информацию и очень давнюю, что сильно опрощают работу

Недостаток единственный в том, что дизайн устаревший

3. Google Translate (<https://translate.google.com.py/?hl=ru>)

* **[Google Translate](https://translate.google.com/)** бесплатный веб-сервис компании Google, предназначенный для перевода отдельных слов, текстов или целых веб-страниц. Уникальность этого переводчика заключается в том, что он работает на базе автономного программного обеспечения, разработанного самой компанией. Google Translate обладает обширным спектром возможностей - он позволяет переводить целые веб-страницы, а также одновременно искать информацию с параллельным переводом на любой другой язык. Благодаря своей быстроте и эффективности, сервис мгновенно осуществляет перевод текстов. Кроме того, Google Translate поддерживает перевод на 103 языка мира.

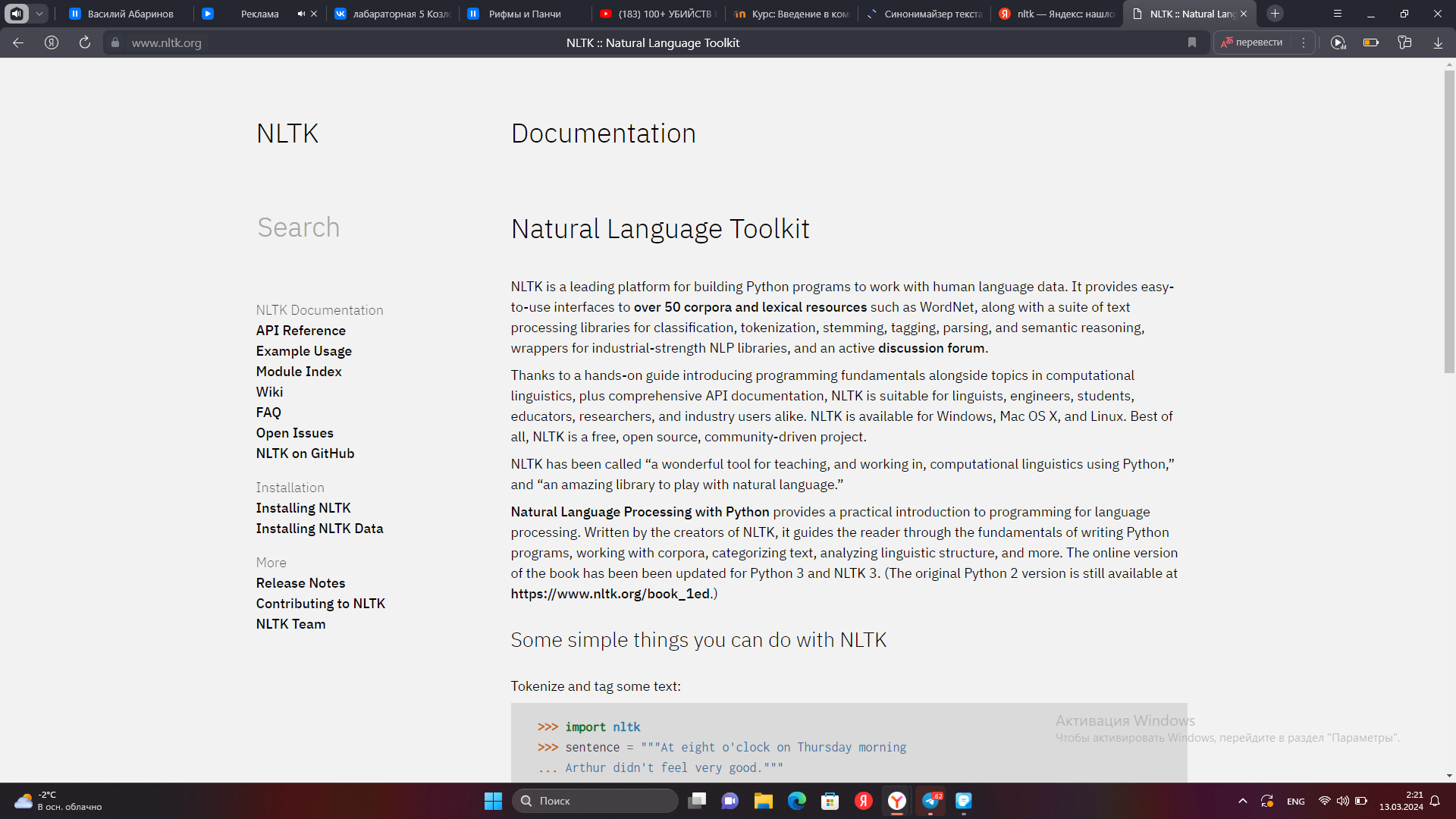
Этот переводчик может переводить не только записанный текст, но и текст на фото сайты, документы, что упрощает пользователю работу. Чтобы перевести текст нужно записать его в левое окно, выбрав язык на котором он написан изначально, и выбрать язык на который нужно перевести. Для перевода текста на фото нужно загрузить это фото на сайт и выбрать язык на который собираетесь переводить.

Достоинство этого ресурса заключается в том, что пользоваться им может даже ребенок. Простое оформление где нельзя запутаться. Множество разных языков. Можно даже прослушать как произноситься переведенный текст(не на всех языках).

Недостаток заключается в том, что перевод не всегда точный.

# 4. [NLTK](https://www.nltk.org/)

**[Natural Language Toolkit](http://www.nltk.org/book/)**Предлагаемый пакет программ и библиотек, разработанный на языке программирования Python, специализируется на символьной и статистической обработке естественного языка. В его состав входят графические представления и примеры данных. Большое внимание уделено документации, которая включает книгу, объясняющую основные концепции, лежащие в основе задач обработки естественного языка, которые можно выполнять с помощью данного пакета.

NLTK является идеальным выбором для студентов, изучающих компьютерную лингвистику или близкие дисциплины, такие как эмпирическая лингвистика, когнитивистика, искусственный интеллект, информационный поиск и машинное обучение. Он успешно применяется в качестве учебного помощника, инструмента для самообучения, а также платформы для создания и прототипирования научно-исследовательских систем. NLTK является свободным программным обеспечением, а проект возглавляет Стивен Бёрд.

NLTK - это платформа на языке Python, предназначенная для работы с естественным языком. Она предлагает простые в использовании интерфейсы для более чем 50 корпусов и лексических ресурсов, включая WordNet, а также набор библиотек для обработки текста, таких как классификация, токенизация, стемминг, тегирование, синтаксический анализ и семантическое рассуждение. Кроме того, NLTK также предоставляет оболочки для мощных библиотек NLP и активно поддерживает дискуссионный форум.

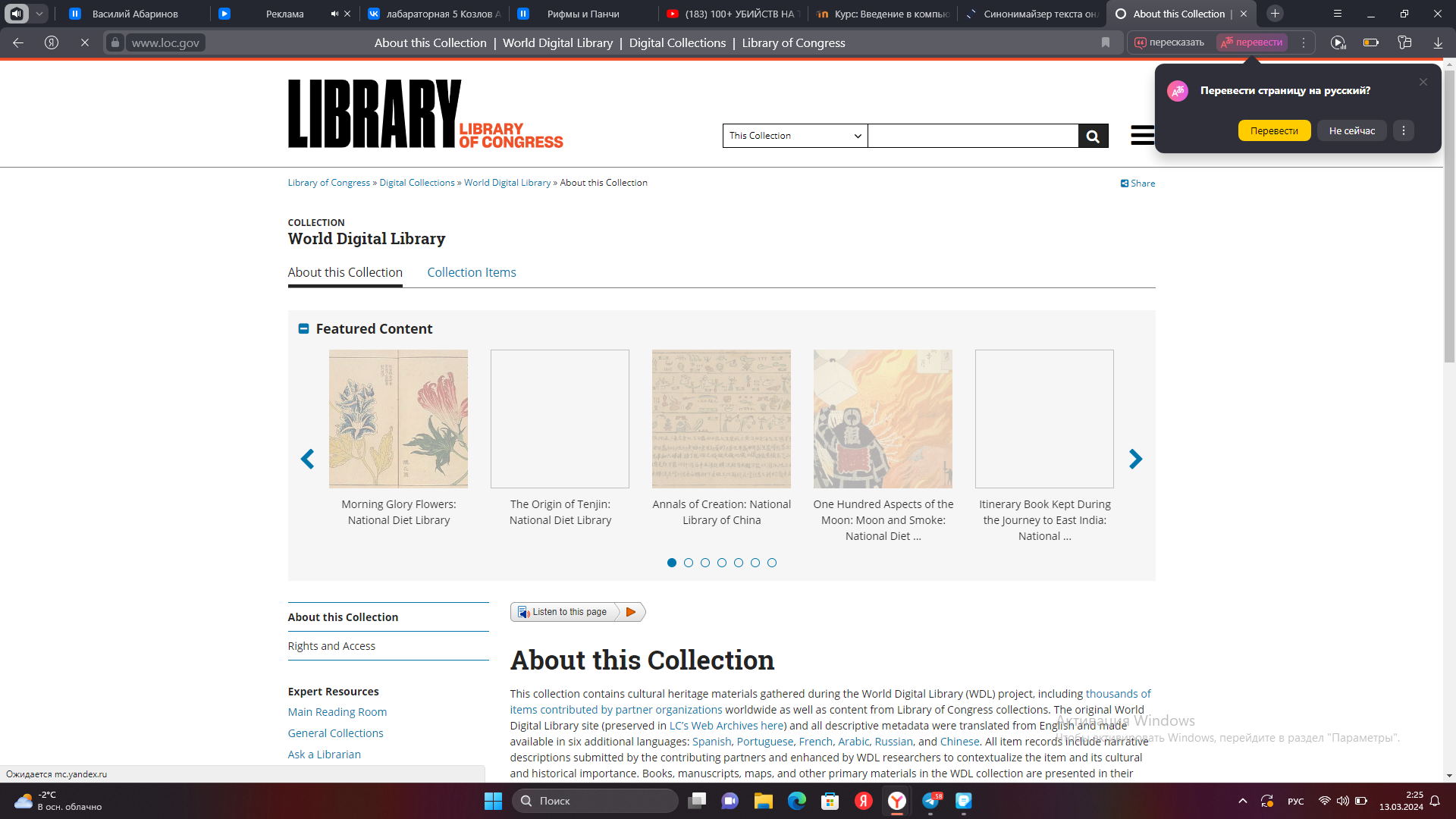
Благодаря своему практическому руководству, которое знакомит читателя с основами программирования на Python, а также с основами компьютерной лингвистики, NLTK подходит для лингвистов, инженеров, студентов, преподавателей, исследователей и отраслевых специалистов. Он доступен для операционных систем Windows, Mac OS X и Linux. Что самое важное, NLTK является бесплатным проектом с открытым исходным кодом и разрабатывается сообществом.

NLTK был признан "замечательным инструментом для преподавания компьютерной лингвистики и работы с ней на Python" и "потрясающей библиотекой для работы с естественным языком".Книга "Обработка естественного языка с использованием Python" является практическим введением в программирование для обработки языка. Она написана создателями NLTK и знакомит читателя с основами написания программ на Python, работой с корпусами, категоризацией текста, анализом лингвистической структуры и многим другим.

Чтобы выполнить поиск, просто нажмите на серый текст "поиск" и в открывшемся окне введите нужный запрос.Достоинства: большая база информации

Недостатки: непонятный в использовании в начале

5. Мировая цифровая библиотека (<https://www.loc.gov/collections/world-digital-library/about-this-collection/>)

Международный проект "World Digital Library" (Всемирная цифровая библиотека) был инициирован библиотекой Конгресса США и с тех пор собрал огромное количество ценных исторических и культурных материалов. В 2007 году к проекту присоединилась Российская национальная библиотека. Он также популярен среди национальных книгохранилищ и архивов разных стран, которые также являются активными участниками. Через сеть Интернет мировая цифровая библиотека предоставляет свободный доступ к огромному объему материалов в многоязычном формате, представляющих культуры различных стран мира.

Достоинства: Огромное количество информации на любые темы

Недостатки: Сайт не полностью на русском

Вопросы:

Единицей корпуса может быть отдельный текстовый документ, предложение, слово или даже символ, в зависимости от выбранной исследователем единицы анализа. В текстовом анализе наиболее часто используется предложение или документ, в то время как в лингвистическом анализе чаще всего используется слово или символ, такие как фонема или морфема.

Отбор текстов для корпуса зависит от целей и задач исследования и может основываться на следующих принципах:

1) Репрезентативность: тексты должны быть репрезентативными для языка или языковой группы, которые исследуются. Например, в Брауновском корпусе были представлены тексты на английском языке из различных жанров и источников, чтобы обеспечить широкое покрытие языковых структур.

2) Разнообразие жанров: в корпусе должны быть представлены тексты различных жанров, таких как публицистика, научные работы, художественная литература и другие, чтобы можно было анализировать специфические языковые особенности в каждом из них.

3) Равномерное покрытие тематик: корпус должен включать тексты на различные тематики, чтобы обеспечить достаточное покрытие языковых вариантов и использования языка в разных сферах.

4) Контролируемые параметры: при отборе текстов может быть учтено также контроль некоторых параметров, например, длина текстов, частота употребления слов, распределение по времени и другие.

Принципы отбора текстов для других корпусов могут быть аналогичными, однако они также могут различаться в зависимости от конкретной задачи исследования. Например, для корпуса новостных текстов может быть важным обеспечить актуальность информации, а для корпуса разговорной речи – представить различные диалекты и региональные варианты языка.

Классификация корпусов может быть различной в зависимости от их характеристик и назначения. Ниже представлены несколько типов корпусов:

1) Исследовательский корпус: создается специально для проведения исследований и анализа в языковой лингвистике и других областях. Такой корпус обычно содержит тексты или данные, собранные для конкретной цели исследования, и может включать в себя различные языки, жанры и стили.

2) Статический корпус: не обновляется или не изменяется с течением времени. Обычно используется для проведения исследований или анализа текстов, зафиксированных на определенный момент времени. Статический корпус может включать старые книги, статьи или письма.

3) Параллельный корпус: содержит параллельные тексты на двух или более языках. Используется для сравнения и анализа переводов, изучения лингвистических явлений и разработки машинного перевода. Параллельные тексты могут быть текстами на исходном языке и их переводами на целевой язык, а также текстами на нескольких языках для межъязыковых сравнительных исследований.

Примером такого исследовательского корпуса является Британский национальный корпус (BNC), который содержит около 100 миллионов словоупотреблений. Он состоит из текстов, собранных из различных источников, таких как разговоры, литература, газеты и другие. Этот корпус репрезентативен по отношению к широкому спектру жанров и стилей текстов, что делает его ценным для лингвистических исследований.

Термины, используемые в корпусной лингвистике, включают:

1) Конкорданс – алфавитно-тематический указатель поисковых запросов в корпусе текстов, который позволяет найти все употребления слова или фразы в контексте.

2) Рандомизация – метод случайного распределения элементов, применяемый в корпусной лингвистике для создания контрольных групп и исключения влияния случайных факторов на результаты исследования.

3) Коллокация – статистически значимая связь между словами или фразами, которая часто встречается в текстах и предполагает их совместное употребление.

4) Подкорпус – часть корпуса текстов, позволяющая проводить более узкую и глубокую аналитику определенной темы или явления.

5) Парсинг – процесс анализа и разбора предложений на составляющие элементы, используемый в корпусной лингвистике для изучения синтаксической структуры текстов.

6) Лемматизация – процесс приведения слов к их базовой форме (лемме), что позволяет учеть все словоформы при анализе корпуса текстов.

7) Корпус-менеджер – программное обеспечение, предназначенное для создания, хранения, поиска и анализа текстовых корпусов, упрощающее работу исследователя в области корпусной лингвистики.

6.

"Corpus Linguistics" на сайте Linguistic Society of America ([https://www.linguisticsociety.org/resource/corpus-linguistics](https://vk.com/away.php?utf=1&to=https%3A%2F%2Fwww.linguisticsociety.org%2Fresource%2Fcorpus-linguistics))  
На данном ресурсе представлены статьи, книги и другие источники по корпусной лингвистике, которые могут помочь исследователям в изучении данной темы.  
  
"The Corpus Linguistics Website" ([http://corpus.byu.edu/](https://vk.com/away.php?utf=1&to=http%3A%2F%2Fcorpus.byu.edu%2F))  
Этот веб-сайт представляет собой базу данных корпусов текстов на различных языках. Здесь можно найти информацию о программном обеспечении для работы с корпусами, а также ссылки на исследования в области корпусной лингвистики.  
  
"Corpora List" на сайте Lancaster University ([http://www.lancaster.ac.uk/fass/projects/corpus/index.htm](https://vk.com/away.php?utf=1&to=http%3A%2F%2Fwww.lancaster.ac.uk%2Ffass%2Fprojects%2Fcorpus%2Findex.htm))  
Этот ресурс содержит список корпусов текстов, которые доступны для исследования, а также информацию о методологии и инструментах корпусной лингвистики.  
  
"Corpus Linguistics" на сайте Oxford Research Encyclopedia of Linguistics ([https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-1074](https://vk.com/away.php?utf=1&to=https%3A%2F%2Foxfordre.com%2Flinguistics%2Fview%2F10.1093%2Facrefore%2F9780199384655.001.0001%2Facrefore-9780199384655-e-1074))  
На данном ресурсе представлен обзор основных концепций и методов корпусной лингвистики, а также актуальные исследования в этой области.

**Лаб№10**

1.

|  |  |  |  |
| --- | --- | --- | --- |
|  | НКРЯ | BNC(English corpora) | COCA |
| Количество вхождений | **39 076** | 1882580 | 49260 |

Значительное проявление лидерства в организации BNC можно объяснить выбором исключительно масштабного корпуса новостей, содержащего более 15 миллиардов слов.

2.

|  |  |  |  |
| --- | --- | --- | --- |
|  | НКРЯ | BNC | COCA |
| 1 | *Меня не трогайте! Я занемог*  *Недугом русским ― чаяньем, надеждой.*  *На зов ее, как птица на манок,*  *Я отозвался с радостью поспешной.* | [Эволюция](https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1214&context=rlj)**[русской](https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1214&context=rlj)**[языковой картины мира и](https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1214&context=rlj)**[культура](https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1214&context=rlj)**[речи](https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1214&context=rlj) | mith &; Nephew countered with its own experts, including a top deputy of the Russian |
| 2 | *«Русским словом»*  *С передовыми Зайцева статьями,*  *Ни (о позор мне!* | <https://www.washingtonpost.com/world/2024/01/16/russia-air-travel-danger-sanctions/>  civilian Russian airplanes experienced at least eight serious mechanical failures | Russian Russian First Year, Part 1... In this introduction to the Russian language, culture and history, students will develop basic skills of speaking, listening, reading, . |
| 3 | *И только вносят жизнь в своем свистящем гаме*  *Лишь русские скворцы, зимующие тут.* | Even so, he wrote, what might, à la rigueur , be acceptable for aRussian Grandmaster | Now what we we talking about? # Although you can say that the Russian mirror performance is poorer than the US ones in some respects, as described above[Moon hoax: why not use telescopes to look at the landers? | Bad] |

3.

|  |  |  |  |
| --- | --- | --- | --- |
| Номер примера | НКРЯ | BYC(English corpora) | COCA |
| 1 | Прилагательное | Прилагательное | Прилагательное |
| 2 | Прилагательное | Прилагательное | Прилагательное |
| 3 | Прилагательное | Прилагательное | Прилагательное |

4.

|  |  |  |  |
| --- | --- | --- | --- |
| № примера | НКРЯ | COCA | BNC |
| 1 | Относится к русской нации | Относится к русской культуре | Относится к политике России |
| 2 | Слово русского языка | Самолеты изготовленные в России | Относится к русской культуре |
| 3 | Птица России | Гроссмеймтер из России | Относится к русскому языку |

5.Взятые слова все являются прилагательными, но лексическое значение у всех разные

6.

Эти инструменты предоставляют возможность осуществлять анализ и проводить исследования в различных языках, диалектах и других лингвистических аспектах. Они оказывают незаменимую помощь в изучении языков.