how to send request to google with js request to find word on gov.pl

**ChatGPT powiedział:**

ChatGPT

To send a request to Google to search for a specific word on a specific site (e.g., gov.pl) using JavaScript, you can use the axios library or the fetch API to make a request to Google’s search engine. Here's an example using axios:

**Using axios:**

javascript

Skopiuj kod

const axios = require('axios');

// Function to perform a Google search

async function searchGoogle(keyword, site) {

const query = `${keyword} site:${site}`;

const googleSearchUrl = `https://www.google.com/search?q=${encodeURIComponent(query)}`;

try {

const response = await axios.get(googleSearchUrl, {

headers: {

'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/91.0.4472.124 Safari/537.36'

}

});

console.log(response.data); // HTML response from Google

} catch (error) {

console.error('Error searching Google:', error.message);

}

}

// Example usage

searchGoogle('climate', 'gov.pl');

**Notes:**

1. **Google Search Results**: The response will be in HTML format since Google does not provide a free JSON API for search results. You can use an HTML parser like cheerio to extract specific data from the response.
2. **Respect Terms of Service**: Scraping Google's search results may violate their Terms of Service. Consider using Google's official Custom Search JSON API for programmatically accessing search results.

**Using Google Custom Search JSON API (Recommended):**

To stay within Google's policies, you can use their Custom Search API. Here's an example:

javascript

Skopiuj kod

const axios = require('axios');

const API\_KEY = 'your\_api\_key'; // Replace with your Google API Key

const CX = 'your\_search\_engine\_id'; // Replace with your Custom Search Engine ID

// Function to perform a Google search using Custom Search API

async function searchGoogleCustom(keyword, site) {

const url = `https://www.googleapis.com/customsearch/v1?q=${encodeURIComponent(keyword)}+site:${encodeURIComponent(site)}&key=${API\_KEY}&cx=${CX}`;

try {

const response = await axios.get(url);

console.log(response.data); // JSON results

} catch (error) {

console.error('Error searching Google:', error.message);

}

}

// Example usage

searchGoogleCustom('climate', 'gov.pl');

**Setup for Custom Search API:**

1. Get a Google API Key.
2. Create a Custom Search Engine and note its ID.
3. Use the API Key and Custom Search Engine ID in the script.

4o

**Powiedziałeś(-aś):**

Error searching Google: Request failed with status code 429

**ChatGPT powiedział:**

ChatGPT

The error **429: Too Many Requests** occurs when the server (in this case, Google) detects too many requests in a short time or considers the request as automated (e.g., scraping). Google actively protects its search engine from bots and scraping attempts.

To resolve this, here are a few alternatives:

**1. Use Google Custom Search API (Recommended)**

Google Custom Search API is designed for programmatic search. Follow these steps:

1. **Set Up API Key:**
   * Visit the Google Cloud Console.
   * Enable the **Custom Search JSON API**.
   * Create an API key.
2. **Set Up Custom Search Engine:**
   * Go to the Custom Search Engine.
   * Create a new search engine and restrict it to gov.pl.
3. **Example Code:** Replace your\_api\_key and your\_search\_engine\_id with your credentials:

javascript

Skopiuj kod

const axios = require('axios');

const API\_KEY = 'your\_api\_key';

const CX = 'your\_search\_engine\_id';

async function searchGoogleCustom(keyword) {

const url = `https://www.googleapis.com/customsearch/v1?q=${encodeURIComponent(keyword)}&key=${API\_KEY}&cx=${CX}`;

try {

const response = await axios.get(url);

console.log(response.data); // JSON results

} catch (error) {

console.error('Error searching Google:', error.message);

}

}

// Example usage

searchGoogleCustom('climate');

**2. Add Delays Between Requests**

If you're using plain axios or fetch to scrape Google, slow down your requests to avoid triggering rate limits:

* Use libraries like [axios-rate-limit](https://www.npmjs.com/package/axios-rate-limit).
* Add a manual delay with setTimeout or similar.

**3. Modify Headers to Mimic Human Behavior**

Google may block requests it deems automated. Add realistic headers:

javascript

Skopiuj kod

const response = await axios.get('https://www.google.com/search?q=climate+site:gov.pl', {

headers: {

'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/91.0.4472.124 Safari/537.36',

'Accept-Language': 'en-US,en;q=0.9',

'Accept-Encoding': 'gzip, deflate, br'

}

});

This may work for a few requests but isn't reliable long-term.

**4. Use a Proxy Service**

Proxy services can rotate IPs to avoid rate limits. Some examples include:

* [ScraperAPI](https://www.scraperapi.com/)
* [Bright Data](https://brightdata.com/)

Modify the request to use a proxy:

javascript

Skopiuj kod

const response = await axios.get('https://www.google.com/search?q=climate+site:gov.pl', {

proxy: {

host: 'proxy-server-address',

port: 8080,

auth: {

username: 'proxy-username',

password: 'proxy-password'

}

}

});

**5. Avoid Scraping Google Directly**

Google’s Terms of Service prohibit scraping. Instead, use:

* **Custom Search API** (as explained above).
* Other search engines with more lenient policies (e.g., Bing, DuckDuckGo).

By using the Google Custom Search API or these alternatives, you can avoid rate limits and stay compliant with Google’s policies.

4o

**Powiedziałeś(-aś):**

give me complete code

**ChatGPT powiedział:**

ChatGPT

Here’s the complete code to use Google’s **Custom Search JSON API** for searching a keyword on gov.pl. This is the recommended and compliant way to programmatically search Google.

**Steps:**

1. **Get an API Key**:
   * Go to the Google Cloud Console.
   * Enable the **Custom Search JSON API** and create an API key.
2. **Create a Custom Search Engine**:
   * Go to Custom Search Engine.
   * Create a new search engine and restrict it to gov.pl.
   * Note down the **Search Engine ID (CX)**.

**Code:**

javascript

Skopiuj kod

const axios = require('axios');

// Replace these with your own API Key and Custom Search Engine ID

const API\_KEY = 'your\_api\_key'; // Replace with your API Key

const CX = 'your\_search\_engine\_id'; // Replace with your Search Engine ID

/\*\*

\* Function to search using Google Custom Search API

\* @param {string} keyword - The keyword to search for

\*/

async function searchGoogle(keyword) {

const url = `https://www.googleapis.com/customsearch/v1?q=${encodeURIComponent(keyword)}&key=${API\_KEY}&cx=${CX}`;

try {

const response = await axios.get(url);

console.log('Search Results:', response.data);

// Return only the items (results) if needed

return response.data.items || [];

} catch (error) {

console.error('Error searching Google:', error.message);

if (error.response) {

console.error('Response Data:', error.response.data);

}

}

}

// Example usage

(async () => {

const keyword = 'climate'; // Replace with your desired keyword

const results = await searchGoogle(keyword);

// Output results

if (results.length > 0) {

results.forEach((item, index) => {

console.log(`${index + 1}. ${item.title} - ${item.link}`);

});

} else {

console.log('No results found.');

}

})();

**Output:**

For a query like climate restricted to gov.pl, the script will print search results as:

perl

Skopiuj kod

1. Title of Result 1 - https://example.gov.pl/link-to-result-1

2. Title of Result 2 - https://example.gov.pl/link-to-result-2

...

**Notes:**

1. **Quota Limits**: The Google Custom Search API allows up to **100 free requests per day**. If you exceed this, you’ll need to enable billing.
2. **Results Filtering**: You can extract and filter specific data like title, link, snippet, etc., from response.data.items.
3. **Error Handling**: The script includes basic error handling. Check the error.response.data for more details if the request fails.

Let me know if you need help setting up the API or customizing this code further!

4o

**Powiedziałeś(-aś):**

i wont search word on gov.pl

**ChatGPT powiedział:**

ChatGPT

If you want to search for a specific word on gov.pl using Google, the correct approach is to construct a query that restricts results to the gov.pl domain. Here’s how you can do it using **Google's Custom Search JSON API**:

**Complete Code for Searching a Word on gov.pl**

javascript

Skopiuj kod

const axios = require('axios');

// Replace these with your own API Key and Custom Search Engine ID

const API\_KEY = 'your\_api\_key'; // Replace with your API Key

const CX = 'your\_search\_engine\_id'; // Replace with your Search Engine ID

/\*\*

\* Function to search for a word on gov.pl using Google Custom Search API

\* @param {string} keyword - The word to search for

\*/

async function searchOnGovPL(keyword) {

// Query restricted to gov.pl domain

const url = `https://www.googleapis.com/customsearch/v1?q=${encodeURIComponent(keyword)}+site:gov.pl&key=${API\_KEY}&cx=${CX}`;

try {

const response = await axios.get(url);

console.log('Search Results:', response.data);

// Return only the items (results) if needed

return response.data.items || [];

} catch (error) {

console.error('Error searching Google:', error.message);

if (error.response) {

console.error('Response Data:', error.response.data);

}

}

}

// Example usage

(async () => {

const keyword = 'climate'; // Replace with your desired word

const results = await searchOnGovPL(keyword);

// Output results

if (results.length > 0) {

results.forEach((item, index) => {

console.log(`${index + 1}. ${item.title} - ${item.link}`);

});

} else {

console.log('No results found.');

}

})();

**How This Works:**

1. **Query Syntax**:
   * The query climate site:gov.pl restricts the search to the gov.pl domain.
   * Replace climate with any word or phrase you want to search for.
2. **API URL**:
   * The API request is constructed with the word and domain restriction:  
     https://www.googleapis.com/customsearch/v1?q=<keyword>+site:gov.pl&key=<API\_KEY>&cx=<CX>
3. **Response**:
   * The response will include search results as JSON data.
   * Each result will have properties like title, link, and snippet.

**Output Example:**

For the word climate, you might see:

ruby

Skopiuj kod

1. Climate Policy on GOV.PL - https://www.gov.pl/web/environment/climate-policy

2. National Climate Strategy - https://www.gov.pl/web/energy/national-strategy

...

**Prerequisites:**

* **API Key**: Follow Google Cloud Console to create and enable the Custom Search JSON API.
* **Search Engine ID (CX)**: Create a Custom Search Engine on Google Custom Search and restrict it to gov.pl.

Let me know if you need further clarification!

4o

**Powiedziałeś(-aś):**

search only domains with gov.pl for example https://legionowo.praca.gov.pl

**ChatGPT powiedział:**

ChatGPT

To search for specific subdomains or URLs under gov.pl—like https://legionowo.praca.gov.pl—you can restrict the search results using **Google's Custom Search JSON API**. Here's how you can modify the query to target only gov.pl subdomains or specific domains like legionowo.praca.gov.pl.

**Full JavaScript Code**

javascript

Skopiuj kod

const axios = require('axios');

// Replace these with your own API Key and Custom Search Engine ID

const API\_KEY = 'your\_api\_key'; // Replace with your API Key

const CX = 'your\_search\_engine\_id'; // Replace with your Search Engine ID

/\*\*

\* Function to search for a word within specific gov.pl domains

\* @param {string} keyword - The keyword to search for

\* @param {string} domain - Specific domain or subdomain to restrict the search (e.g., "legionowo.praca.gov.pl")

\*/

async function searchOnSpecificGovPL(keyword, domain = 'gov.pl') {

// Query restricted to the specific domain

const url = `https://www.googleapis.com/customsearch/v1?q=${encodeURIComponent(

keyword

)}+site:${domain}&key=${API\_KEY}&cx=${CX}`;

try {

const response = await axios.get(url);

console.log('Search Results:', response.data);

// Return only the items (results) if available

return response.data.items || [];

} catch (error) {

console.error('Error searching Google:', error.message);

if (error.response) {

console.error('Response Data:', error.response.data);

}

}

}

// Example usage

(async () => {

const keyword = 'job'; // Replace with your search term

const domain = 'legionowo.praca.gov.pl'; // Replace with your desired domain or use 'gov.pl'

const results = await searchOnSpecificGovPL(keyword, domain);

// Output results

if (results.length > 0) {

results.forEach((item, index) => {

console.log(`${index + 1}. ${item.title} - ${item.link}`);

});

} else {

console.log('No results found.');

}

})();

**How It Works:**

1. **Query Syntax**:
   * The query is built as:

php

Skopiuj kod

<keyword> site:<domain>

For example: job site:legionowo.praca.gov.pl.

1. **Restrict to Subdomain**:
   * The site:<domain> operator in the query ensures that results come only from the specified subdomain (e.g., legionowo.praca.gov.pl).
2. **Response**:
   * Each result in response.data.items contains:
     + title: Title of the result
     + link: URL of the result
     + snippet: Short description of the result
3. **Fallback to General Domain**:
   * If you don't provide a specific domain, it defaults to gov.pl (searching all subdomains under gov.pl).

**Output Example:**

For the keyword job and domain legionowo.praca.gov.pl, you might see:

arduino

Skopiuj kod

1. Job Offers - https://legionowo.praca.gov.pl/job-offers

2. Local Employment Office - https://legionowo.praca.gov.pl/contact

...

**Notes:**

* **Domain Restriction**:
  + Replace legionowo.praca.gov.pl with any specific subdomain or leave it as gov.pl for a broader search.
* **Custom Search Engine**:
  + Ensure your Custom Search Engine (CX) is configured to allow searches on gov.pl domains.
* **Quota**:
  + The API allows up to 100 free queries per day. For more, enable billing in the Google Cloud Console.

Let me know if you need help configuring the API or understanding the results!

4o

**Powiedziałeś(-aś):**

dont use. const domain = 'legionowo.praca.gov.pl'; but find on domains whitch hace gov.pl

**ChatGPT powiedział:**

ChatGPT

To search all subdomains under gov.pl without hardcoding a specific domain like legionowo.praca.gov.pl, you can still use the **site** search parameter (site:gov.pl). The following JavaScript code will search for your keyword across all domains and subdomains ending in gov.pl.

**Complete Code**

javascript

Skopiuj kod

const axios = require('axios');

// Replace these with your own API Key and Custom Search Engine ID

const API\_KEY = 'your\_api\_key'; // Replace with your Google API Key

const CX = 'your\_search\_engine\_id'; // Replace with your Custom Search Engine ID

/\*\*

\* Function to search for a keyword across all gov.pl domains

\* @param {string} keyword - The word to search for

\*/

async function searchAcrossGovPL(keyword) {

// Query restricted to all domains under gov.pl

const url = `https://www.googleapis.com/customsearch/v1?q=${encodeURIComponent(

keyword

)}+site:gov.pl&key=${API\_KEY}&cx=${CX}`;

try {

const response = await axios.get(url);

console.log('Search Results:', response.data);

// Return only the items (results) if available

return response.data.items || [];

} catch (error) {

console.error('Error searching Google:', error.message);

if (error.response) {

console.error('Response Data:', error.response.data);

}

}

}

// Example usage

(async () => {

const keyword = 'job'; // Replace with your search term

const results = await searchAcrossGovPL(keyword);

// Output results

if (results.length > 0) {

results.forEach((item, index) => {

console.log(`${index + 1}. ${item.title} - ${item.link}`);

});

} else {

console.log('No results found.');

}

})();

**Key Changes:**

1. **Site Query**:
   * The query string site:gov.pl restricts the search to any domain or subdomain ending with .gov.pl.
   * For example:
     + www.praca.gov.pl
     + warszawa.praca.gov.pl
     + legionowo.praca.gov.pl
2. **Dynamic Coverage**:
   * This approach doesn't hardcode specific domains but dynamically covers all subdomains under gov.pl.
3. **Result Parsing**:
   * Results are parsed dynamically from the Google Custom Search JSON API and returned as a list of title and link.

**Example Output:**

For the keyword job on all gov.pl subdomains, results might look like this:

arduino

Skopiuj kod

1. Job Offers - https://warszawa.praca.gov.pl/job-offers

2. Local Employment Office - https://legionowo.praca.gov.pl/contact

3. Employment Portal - https://praca.gov.pl/job-portal

**Notes:**

* **Custom Search Engine Setup**:
  + Ensure your Custom Search Engine (CX) is set up to include gov.pl domains.
* **API Key**:
  + Obtain your API key and CX from Google Cloud Console.
* **Quota**:
  + Each request counts against your Google Custom Search API quota (100 free requests/day).

Let me know if you need assistance configuring your API or running this code!

4o

**Powiedziałeś(-aś):**

is it paid from google/

**ChatGPT powiedział:**

ChatGPT

Yes, using the **Google Custom Search JSON API** can incur charges if you exceed the free tier limits. Here's an overview of the cost structure as of now:

**Google Custom Search JSON API Pricing**

1. **Free Tier**:
   * **Requests**: Up to 100 queries per day.
   * **Cost**: Free.
2. **Paid Tier**:
   * **Additional Queries**: $5 per 1,000 queries beyond the free 100/day.
   * **Billing**: Billed monthly via your Google Cloud account.

**Key Notes:**

* **Free Usage**:
  + For small-scale personal or development projects, the free tier is often sufficient.
* **Paid Usage**:
  + For larger-scale or production-grade applications, charges apply after the free tier.

**Alternatives to Avoid Payment:**

1. **Manual Web Scraping**:
   * You can scrape gov.pl search results using libraries like **Puppeteer** or **Cheerio**, but scraping Google directly is against their terms of service and can lead to IP blocking.
2. **Use Public APIs**:
   * Check if gov.pl or its subdomains provide a public API for searching their data. This is a more ethical and reliable option than scraping Google.
3. **Custom Search Engine**:
   * Set up a free **Custom Search Engine (CSE)** on Google that indexes only gov.pl domains. You can use the free interface (without API calls) if suitable for your project.

Would you like help with an alternative approach, such as using a public API or setting up a Custom Search Engine? Let me know!