A Novel Approach to Scraping Websites

ROB OUSBHEY
VP, Product, Moz

@robousbey
Where you go to buy software.

You can choose the right software and services for your business based on 1,086,800+ authentic, timely reviews from real users.

Enter product name, software category, service name, ...

Search

Best Products 2020
Trending Products

Most Popular Software

Asana

Smartsheet

monday.com

rob.ousbey@moz.com
@RobOusbey
Warning: Code Ahead
Paths forward:

A: Just use the code that I share today

B: Buddy up with a real developer

C: Lean some JavaScript
Learn JavaScript?

Tips for becoming an effective Javascript hacker

- Learn a little bit of JavaScript, and a lot of jQuery
- Start with existing code, and edit it to understand how it works
- Google everything - it's all been solved before
- Write code slowly and check / debug it a bit at a time
- Ignore everything about code quality, elegance, efficiency and scalability
- If it works, then it works

If you're learning JavaScript, you've likely heard people tell you how important it is to learn the fundamentals. But what are they? And where do you learn them?

Here's a list of JavaScript fundamentals and my favorite free resources for learning them.

12:29 PM - Jun 20, 2020 - Twitter Web App

2.5K Retweets and comments 8.6K Likes
A Novel Approach to Scraping Websites
A Novel Approach to Scraping Websites
A Novel Approach to Scraping Websites
Insert your code into any website

(Using Google Chrome)
Option 1: Write / paste code directly into the Console

Press F12 to open Chrome DevTools & click on the ‘Console’ tab
Color Of Change helps you do something real about injustice.

We design campaigns powerful enough to end practices that unfairly hold Black people back, and champion solutions that move us all forward. Until justice is real.
$(selector)

Finds an element on the current page, using the typical CSS selectors that you're used to.
$(selector)

Finds an element on the current page, using the typical CSS selectors that you're used to.
Option 2: Store code in a ‘JavaScript bookmarklet’

- Storing code in a bookmarklet lets you click to run it anytime you want
- Begin the URL with `javascript:` and then add your code
- Lines of code should be separated by semi-colons
Option 3: Store code in an online JavaScript file

- Host your code somewhere accessible online (including Dropbox, etc)
- Create a bookmarklet (see option 2) that does nothing but import this code.
- Include it on the page by creating a new script element, using code like:

```javascript
javascript:(function(){
    document.body.appendChild(document.createElement('script')).src='https://ousbey.com/mycode.js';
})();
```
Option 3: Store code in an online JavaScript file

Host your code somewhere accessible online (including Dropbox, etc)
Create a bookmarklet (see option 2) that does nothing but import this code.
Include it on the page by creating a new script element, using code like:

```javascript
javascript:(function(){
    document.body.appendChild(document.createElement('script')).
    src='https://ousbey.com/mycode.js?x='+Math.random();})();
```

PRO TIP!
Chapter 1:
Some basic scraping
Pseudo Code

For each one of multiple products, we will:

• Search for the product name
Pseudo Code

For each one of multiple products, we will:

- Search for the product name
- Get the URL from the top result
Pseudo Code

For each one of multiple products, we will:

- Search for the product name
- Get the URL from the top result
- Go to that URL, and scrape:
  - the overall score
Pseudo Code

For each one of multiple products, we will:

• Search for the product name
• Get the URL from the top result
• Go to that URL, and scrape:
  - the overall score
  - the detailed ratings
Pseudo Code

For each one of multiple products, we will:

• Search for the product name
• Get the URL from the top result
• Go to that URL, and scrape:
  - the overall score
  - the detailed ratings
• Put all that data in a table of some kind
Your search: **zendesk chat** returned 6 results

### Products (4)

**Zendesk Chat**

★★★★★ (399) 4.3 out of 5

Zendesk Chat's live chat solution helps businesses increase sales conversion by engaging important leads on their websites.

[Reviews] [Alternatives] [Other Live Chat Software]

**Pythia for Zendesk**

★★★★★ (9) 5.0 out of 5

Pythia is a one-stop-shop of affordable AI tools for scaling the productivity of your Zendesk. - Advanced Macros helps to utilize macros in Zendesk Support. - Advanced Shortcuts helps to utilize shrtcuts in Zendesk Chat. - Macros Reporting provide...

[Reviews] [Alternatives] [Other Productivity Bots Software]

**LiveChat**

★★★★☆ (689) 4.4 out of 5

Premium live chat and help desk software for business websites. E-commerce and customer service software with built-in ticketing system.
Products

Zendesk Chat

⭐⭐⭐⭐⭐ (399) 4.3 out of 5
Zendesk Chat's live chat solution helps businesses increase sales conversion by engaging important leads on their websites.

Reviews | Alternatives | Other Live Chat Software

Pythia for Zendesk

⭐⭐⭐⭐⭐ (5) 5.0 out of 5
Pythia is a one-stop-shop of affordable AI tools for scaling the productivity of your Zendesk. - Advanced Macros helps to utilize macros in Zendesk Support. - Advanced Shortcuts helps to utilize shprtcuts in Zendesk Chat. - Macros Reporting provide...

Reviews | Alternatives | Other Productivity Bots Software

LiveChat

⭐⭐⭐⭐⭐ (689) 4.4 out of 5
Premium live chat and help desk software for business websites. E-commerce and customer service software with built-in ticketing system.
Get data from a URL; Parse the response

```javascript
searchQuery = "zendesk";

$.get( "/search/products?query="+ searchQuery, function( data ) {

    var $page = $(data);
    console.log( $page.find("h1").text() );

    topListing = $( $page.find(".product-listing")[0] );
    productData = {
        url: topListing.find(".product-listing__title a").attr("href"),
        name: topListing.find(".product-listing__product-name").text(),
        rating: topListing.find(".product-listing__star-rating .fw-semibold").first().text()
    }

    console.log(productData);

});
```
Get data from a URL; Parse the response

searchQuery = "zendesk";

$.get( "/search/products?query=" + searchQuery, function( data ) {
  var $page = $(data);
  console.log( $page.find("h1").text() );

  topListing = $( $page.find(".product-listing")[0] );
  productData = {
    url: topListing.find(".product-listing__title a").attr("href"),
    name: topListing.find(".product-listing__product-name").text(),
    rating: topListing.find(".product-listing__star-rating .fw-semibold").first().text()
  };

  console.log(productData);
});

$.get(url, function) Fetch a page from elsewhere on the site, and process the result
Get data from a URL; Parse the response

```javascript
searchQuery = "zendesk";

$.get( "/search/products?query=" + searchQuery,
    function(data) {
        var $page = $(data);
        console.log( $page.find("h1").text() );

        topListing = $( $page.find(".product-listing")[0] );
        productData = {
            url: topListing.find(".product-listing__title__a").attr("href"),
            name: topListing.find(".product-listing__product-name").text(),
            rating: topListing.find(".product-listing__star-rating .fw-semibold").first().text()
        }
        console.log( productData );
    });
```

**.text()**
- Grab the text from a selected element

**.find(selector)**
- Parse a page, and find the one or more elements that match the CSS selector
```javascript
searchQuery = "zendesk";

$.get( "/search/products?query=" + searchQuery, function( data ) {

    var $page = $(data);
    console.log( $page.find("h1").text() );

    topListing = $( $page.find(".product-listing")[0] );
    productData = {
        url: topListing.find(".product-listing__title a").attr("href"),
        name: topListing.find(".product-listing__product-name").text(),
        rating: topListing.find(".product-listing__star-rating .fw-semibold").first().text()
    }

    console.log(productData);
});
```

Your search: zendesk returned 67 results

```json
{"url": "/products/zendesk-support/reviews", "name": "Zendesk Support", "rating": "4.2 "}
```
Your search: zendesk returned 67 results

- {url: "/products/zendesk-support/reviews", name: "Zendesk Support", rating: "4.2"}
 zendesk chat

Product Information  Reviews  Pricing  Features

with customers in real-time and let me discuss about product features.

What do you dislike?
It does not enable you to edit your reply once it is delivered and another dislike in this application is that it is quite expensive as compared to other alternatives. They don't provide network monitoring and there is no work area morning, so need more updates to solve these issues for better experience.

Show More

 Zendesk Chat User Ratings

Ease of Use
8.9
Live Chat Average: 8.8

Quality of Support
8.4
Live Chat Average: 8.6

Ease of Setup
8.6
Live Chat Average: 8.7

Compare Zendesk Chat to similar products

Do you work for Zendesk Chat?
Update this Profile

Very useful program with after work ticketing system"

What do you like best?
I think the best aspect of Zendesk Chat was the automated tickets that populate after you close out of a chat. This allows for after chat notes which will be useful when a customer chats back in or to organize and collect data on trends. The program is also very easy to use and efficiently speak with customers in realtime, as well as speaking to multiple customers at a time quite easily. The notification system is fantastic in letting you know when a customer has responded, especially if you have closed out their window.

What do you dislike?

Get data from a URL; Parse the response

```
$.get(productData.url, function( pageContent ) {
    var $productPage = $(pageContent);

    ratingDetails = {};

    ratingDetails.use = $productPage.find('.cell.small-7 div:contains("Ease of Use")').closest('.grid-x').find('.charts--doughnut__reviews').text();

    ratingDetails.support = $productPage.find('.cell.small-7 div:contains("Quality of Support")').closest('.grid-x').find('.charts--doughnut__reviews').text();

    ratingDetails.setup = $productPage.find('.cell.small-7 div:contains("Ease of Setup")').closest('.grid-x').find('.charts--doughnut__reviews').text();
});
```
$.get(productData.url, function( pageContent ) {

  var $productPage = $(pageContent);

  ratingDetails = {};

  ratingDetails.use = $productPage.find('.cell.small-7
   div:contains("Ease of Use")').closest('.grid-x')
   .find('.charts--doughnut__reviews').text();

  ratingDetails.support = $productPage.find('.cell.small-7
   div:contains("Quality of Support")').closest('.grid-x')
   .find('.charts--doughnut__reviews').text();

  ratingDetails.setup = $productPage.find('.cell.small-7
   div:contains("Ease of Setup")').closest('.grid-x')
   .find('.charts--doughnut__reviews').text();
});
$.get(productData.url, function( pageContent ) {

    var $productPage = $(pageContent);

    ratingDetails = {};

    ratingDetails.use = $productPage.find('.cell.small-7
div:contains("Ease of Use")').closest('.grid-x').
    .find('.charts--doughnut__reviews').text();

    ratingDetails.support = $productPage.find('.cell.small-7
div:contains("Quality of Support")').closest('.grid-x').
    .find('.charts--doughnut__reviews').text();

    ratingDetails.setup = $productPage.find('.cell.small-7
div:contains("Ease of Setup")').closest('.grid-x').
    .find('.charts--doughnut__reviews').text();

});
> ratingDetails

← ▷ {use: "8.9", support: "8.4", setup: "8.6"}

> │
Writing Data

```javascript
$.html('#container').html('<div id="output"></div>');

$.append('#output').append('<span class="product" data-id="zendesk"></span>');

$.text('span[data-id="zendesk"]').text('4.3');
```

- `.text(data)`  
  Replace the text of an element

- `.html(data)`  
  Replace the HTML of an element

- `.append(data)`  
  Add html on to the end of an element's content
Writing data to a table

```html
$("html").html(`
<html>
<head></head>
<body>

<textarea id="input"></textarea>
<button id="start">Go</button>

<table id="output">
  <tr>
    <th>Product</th>
    <th>Reviews</th>
    <th>Score</th>
    <th>Ease of Use</th>
    <th>Support</th>
    <th>Ease of Setup</th>
  </tr>
</table>

</body>
</html>`);
```
Product Reviews Score Ease of Use Support Ease of Setup
function populateTable()
{
    productList = $('#input').val().split(/\r?\n/);

    $.each( productList, function( i, inputName ){
        inputName = inputName.trim().toLowerCase();

        $('#output').append('<tr class="product" data-name="' + inputName + '", data-processed="false"><td class="name">' + inputName + '</td><td class="count"></td><td class="rating"></td><td class="rating_use"></td><td class="rating_support"></td><td class="rating_setup"></td></tr>');
    });
}
Adding interactivity

```
$('[#start]').click(function() {
    populateTable();
    startCrawl();
});
```

.click( function)

Specifies code to be run when you click the selected element(s)
<table>
<thead>
<tr>
<th>Product</th>
<th>Reviews Score</th>
<th>Ease of Use</th>
<th>Support</th>
<th>Ease of Setup</th>
</tr>
</thead>
<tbody>
<tr>
<td>xero</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>freshbooks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>quickbooks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>microsoft dynamics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>netsuite</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
function startCrawl(){

  $("#output tr").each(function( index ) {
    crawlSearchResults( $(this) );
  });
}

Loops
Where you go to buy software.

You can choose the right software and services for your business based on 1,086,800+ authentic, timely reviews from real users.

Enter product name, software category, service name, ...

Search

Best Products 2020
Trending Products

Most Popular Software

Maximize productivity and improve collaboration with Aana.

Smartsheet provides a simple, intuitive interface for teams to plan, track, and report on work.

See all Project Management Software
# G2 Scraper

Enter product names, each on a new line

<table>
<thead>
<tr>
<th>Product</th>
<th>Reviews</th>
<th>Score</th>
<th>Ease of Use</th>
<th>Support</th>
<th>Ease of Setup</th>
</tr>
</thead>
</table>
G2 Scraper

- xero
- freshbooks
- quickbooks
- microsoft dynamics
- netsuite

Go

<table>
<thead>
<tr>
<th>Product</th>
<th>Reviews</th>
<th>Score</th>
<th>Ease of Use</th>
<th>Support</th>
<th>Ease of Setup</th>
</tr>
</thead>
</table>

# G2 Scraper

Enter product names, each on a new line

<table>
<thead>
<tr>
<th>Product</th>
<th>Reviews</th>
<th>Score</th>
<th>Ease of Use</th>
<th>Support</th>
<th>Ease of Setup</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xero</td>
<td>333</td>
<td>4.3</td>
<td>8.7</td>
<td>7.5</td>
<td>8.1</td>
</tr>
<tr>
<td>FreshBooks</td>
<td>571</td>
<td>4.6</td>
<td>9.2</td>
<td>9.3</td>
<td>9.3</td>
</tr>
<tr>
<td>QuickBooks Online</td>
<td>2326</td>
<td>4.0</td>
<td>8.1</td>
<td>7.5</td>
<td>7.9</td>
</tr>
<tr>
<td>Microsoft Dynamics 365 Sales</td>
<td>1428</td>
<td>3.7</td>
<td>7.4</td>
<td>7.4</td>
<td>7.1</td>
</tr>
<tr>
<td>NetSuite</td>
<td>1106</td>
<td>3.9</td>
<td>7.6</td>
<td>7.0</td>
<td>6.5</td>
</tr>
</tbody>
</table>
Use the G2 Scraper for yourself!

Inspect the code for yourself, or install the bookmarklet from here:

ousbey.com / mozcon
Who else can we scrape?
Chapter 2: When the scraper becomes the scrapee
<table>
<thead>
<tr>
<th>URL</th>
<th>Status Code</th>
<th>Issues</th>
<th>Page Authority</th>
<th>Crawl Depth</th>
<th>Analyze</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distilled: Smarter Online Marketing</td>
<td>200</td>
<td>0</td>
<td>56</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Distilled</td>
<td>SearchLove Conference</td>
<td>200</td>
<td>0</td>
<td>51</td>
<td>2</td>
</tr>
<tr>
<td>Online Marketing &amp; Sales</td>
<td>200</td>
<td>3</td>
<td>50</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Distilled</td>
<td>SearchLove Conference</td>
<td>200</td>
<td>0</td>
<td>50</td>
<td>2</td>
</tr>
<tr>
<td>Find (Almost) Anybody's Email Address</td>
<td>200</td>
<td>2</td>
<td>48</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Join Our Team</td>
<td>Distilled</td>
<td>200</td>
<td>2</td>
<td>48</td>
<td>2</td>
</tr>
<tr>
<td>Online Marketing Blog</td>
<td>200</td>
<td>5</td>
<td>48</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Building Your Mobile Site - The Distilled Best Practice Guide</td>
<td>200</td>
<td>2</td>
<td>48</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>The ImportXML Guide</td>
<td>200</td>
<td>1</td>
<td>47</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4 Types of Content Executive</td>
<td>200</td>
<td>1</td>
<td>47</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Content Marketing: How to Create Your Content Strategy</td>
<td>Distilled</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Let’s answer the question:

How many pages does Google have indexed from each folder on the site?
Pseudo Code

For the site in question, we will:

• Run a “site:domain.com” search
Pseudo Code

For the site in question, we will:

- Run a “site:domain.com” search
- Make a list of all the folders we find

Moz.com:

/learn
/blog
/community
Pseudo Code

For the site in question, we will:

• Run a “site:domain.com” search
• Make a list of all the folders we find
• Do another site: search for each folder to count the number of indexed pages
Pseudo Code

For the site in question, we will:

• Run a “site:domain.com” search
• Make a list of all the folders we find
• Do another site: search for each folder to count the number of indexed pages
• Record the results

Moz.com:

/learn – 447 pages
/blog – 7,900 pages
/community – 54,000 pages
Pseudo Code

For the site in question, we will:

• Run a “site:domain.com” search
• Make a list of all the folders we find
• Do another site: search for each folder to count the number of indexed pages
• Record the results
• Do another site: search - excluding the folder you already found - to discover more folders and continue the process
Replace page with input / output UI

```javascript
$("html").html(`
<html>
  <head></head>
  <body>
    <textarea id="input"></textarea>
    <table id="output"><tr>
      <th>Folder</th>
      <th>Indexed Pages</th>
    </tr></table>
    <button id="collect_serps">Get more folders</button>
    <button id="get_counts">Count indexed Pages</button>
  </body>
</html>`);
```
Folder Indexed Pages

Get more folders  Count indexed Pages
Google Scraper - Site Structure

www.example.com

Crawl Site

Folder  Indexed Pages
Find an initial list of folders on the site

```javascript
var domain = ($('##input').val());

var query = 'site:' + domain;

$.get( '/search?q=' + query, function( data ) {
    var $page = $(data);

    $page.find('a[href^="https://"+domain+]').each(function(index){
        var url = $(this).attr('href') + '/';
        var slug = url.split('://')[1].split('/')[1].toLowerCase();

        $('#output').append('<tr class="folder" data-name="'+slug+'" data-processed="false"><td class="slug">'+slug+'</td><td class="count"></td></tr>');
    });
});
```
Google Scraper - Site Structure

moz.com

Crawl Site

Folder  Indexed Pages
Count the indexed pages in any given folder

```javascript
function getFolderCount(slug){
  $.get( "/search?q=site%3A" + domain+"/"+slug+"/", function( data ){
    var $page = $(data);
    count = $page.find('#result-stats').text().replace('About ', '').split(' ')[0]
    $('#output tr[data-name="'+slug+'"] td.count').text(count);
  });
}
```
About 7,410 results (0.24 seconds)
<div>
  <div id="result-stats">
    "About 7,410 results"
    <nobr>(0.24 seconds)&nbsp;</nobr>
  </div>
</div>
Count the indexed pages in any given folder

```javascript
function getFolderCount(slug){

  $.get( "/search?q=site%3A"+ domain+"/"+slug+"/", function( data ){

    var $page = $(data);

    count = $page.find('#result-stats').text().replace('About ', '').split(' ')[0]

    $('#output tr[data-name="'+slug+'"] td.count').text(count);

  });
}
```
Limit the scraping speed

```javascript
var crawlDelay = 1000; // timing in milliseconds

function getAllCounts()
{
    scheduledTimeFromNow = 0;

    $('#output tr').each(function(index)
    {
        scheduledTimeFromNow += crawlDelay;
        getFolderCountWithDelay($(this).attr('data-name'), scheduledTimeFromNow);
    });

    function getFolderCountWithDelay(slug, delay)
    {
        setTimeout(function(){getFolderCount(slug);}, delay);
    }
```

`setTimeout(function, delay)`
Schedules the function to be run in a certain number of milliseconds in the future
Use this Google Scraper for yourself!

Inspect the code for yourself, or install the bookmarklet from here:

ousbey.com / mozcon
Chapter 3:
Multi-site Scraping
Can we grab some SEO data?
Britney Muller - SME 12:34 PM
Could you crawl this: https://web.dev/measure/

The Lighthouse API is the worst

Rob Ousbey 12:35 PM
Oh, PERFECT

Will take a look

Britney Muller - SME 12:35 PM
THANK YOU!!!! Ah, that would be so cool!!!

Rob Ousbey 7:47 PM
All righty - I'm diving in...

Rob Ousbey 9:48 PM
OK, I tried something fun with it and I got it working.

Might not be what you're expecting, but I'm pretty pleased with the last couple hours of work 😊
Measure

See how well your website performs. Then, get tips to improve your user experience. Sign in to track your progress over time.

Enter a web page URL

Performance 0 Accessibility 0 Best Practices 0 SEO 0

Score scale: 0-49 50-89 90-100

1. Run tests on your site
   Enter your site's URL to see how well it performs across all audits.

2. Look at what matters
   See your site's performance across the areas you care about.

3. Get tips for improving
   Each test comes with helpful steps to improve your site's results.
<table>
<thead>
<tr>
<th>Name</th>
<th>Status</th>
<th>Type</th>
<th>Initiator</th>
<th>Size</th>
<th>Time</th>
<th>Waterfall</th>
</tr>
</thead>
<tbody>
<tr>
<td>newaudit</td>
<td>200</td>
<td>text/plain</td>
<td>Other</td>
<td>0 B</td>
<td>143 ms</td>
<td></td>
</tr>
<tr>
<td>collect</td>
<td>200</td>
<td>gif</td>
<td>analytics.js:36</td>
<td>45 ms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Service...)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>collect</td>
<td>200</td>
<td>fetch</td>
<td>fetchWrapper.js</td>
<td>80 B</td>
<td>41 ms</td>
<td></td>
</tr>
<tr>
<td>newaudit</td>
<td>201</td>
<td>fetch</td>
<td>lighthouse-service</td>
<td>235 kB</td>
<td>12.55 s</td>
<td></td>
</tr>
<tr>
<td>log?format=json_proto</td>
<td>200</td>
<td>fetch</td>
<td>cc_service.ts:440</td>
<td>18 ms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>reports?url=https%3A%2F%2F...</td>
<td>200</td>
<td>fetch</td>
<td>lighthouse-service</td>
<td>84.6 kB</td>
<td>1.55 s</td>
<td></td>
</tr>
<tr>
<td>log?format=json_proto</td>
<td>200</td>
<td>fetch</td>
<td>cc_service.ts:90</td>
<td>562 B</td>
<td>20 ms</td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Headers</td>
<td>Preview</td>
<td>Response</td>
<td>Initiator</td>
<td>Timing</td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>---------</td>
<td>----------</td>
<td>-----------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>newaudit</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>collect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log?format=json_proto</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reports?url=https%3A%2F%2F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Request URL:** https://lighthouse-dot-webdotdevsite.appspot.com//lh/newaudit

**Request Method:** POST

**Status Code:** 201

**Remote Address:** 172.217.3.180:443

**Referrer Policy:** no-referrer-when-downgrade

**Response Headers (12)**

**Request Headers (15)**

**Request Payload**

```javascript
//{url: "https://www.ousbey.com", replace: true, save: false}
```

7 requests | 320 KB transferred
<table>
<thead>
<tr>
<th>Name</th>
<th>Headers</th>
<th>Preview</th>
<th>Response</th>
<th>Initiator</th>
<th>Timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>newaudit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>collect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>collect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>newaudit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

```json
{1hrSlim: [{id: "performance", title: "Performance", score: 0.61},...],
  auditedOn: "2020-06-24T20:46:05.164Z"
}
```

- `-0`: {id: "performance", title: "Performance", score: 0.61}
- `-1`: {id: "accessibility", title: "Accessibility",...}
- `-2`: {id: "best-practices", title: "Best Practices", score: 0.85}
- `-3`: {id: "seo", title: "SEO"}

Description: These checks ensure that your page is optimized for search engines.

id: "seo"

manualDescription: "Run these additional validators on your site to check..."
Take a 30 day free trial, at moz.com/pro
# SERP Analysis

<table>
<thead>
<tr>
<th>Rank</th>
<th>Title &amp; URL</th>
<th>PA Linking RDs to Page</th>
<th>DA Linking RDs to Root Domain</th>
<th>Page Optimization Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>SEO Agency</td>
<td>Industry-Recommended Online SEO Agency ... <a href="http://www.webfx.com/seo-agency.html">www.webfx.com/seo-agency.html</a></td>
<td>43</td>
<td>64</td>
</tr>
<tr>
<td>#2</td>
<td>SEO Agency</td>
<td>Top SEO Agency</td>
<td>SEO Services - Actuate Media <a href="http://www.actuatemedia.com/seo-agency/">www.actuatemedia.com/seo-agency/</a></td>
<td>27</td>
</tr>
<tr>
<td>#3</td>
<td>SEO Agency</td>
<td>One Click SEO Is Your Full Service SEO ... oneclickseo.agency</td>
<td>35</td>
<td>31</td>
</tr>
<tr>
<td>#4</td>
<td>SEO Company</td>
<td>Search Engine Optimization Firm - SEO ... <a href="http://www.seo.com">www.seo.com</a></td>
<td>61</td>
<td>52</td>
</tr>
<tr>
<td>#5</td>
<td>Best SEO Companies &amp; Services - 2020 Reviews</td>
<td>Clutch.co clutch.co/seo-frms</td>
<td>52</td>
<td>62</td>
</tr>
<tr>
<td>#6</td>
<td>SEO Agency</td>
<td>World-Class SEO Company</td>
<td>Directive directiveconsulting.com/services/seo/</td>
<td>37</td>
</tr>
<tr>
<td>#7</td>
<td>SEO Company</td>
<td>Thrive Internet Marketing Agency thriveagency.com/...-marketing-services/search-engine-optimization-seo/</td>
<td>36</td>
<td>43</td>
</tr>
<tr>
<td>#8</td>
<td>SEO Agency</td>
<td>HighMark SEO Digital</td>
<td>Engineered to ... <a href="http://www.highmarkseo.com">www.highmarkseo.com</a></td>
<td>37</td>
</tr>
<tr>
<td>#9</td>
<td>#1 Los Angeles SEO Company</td>
<td>Search Engine Optimization coalitiontechnologies.com/seo-search-engine-optimization</td>
<td>40</td>
<td>38</td>
</tr>
<tr>
<td>Rank</td>
<td>Title &amp; URL</td>
<td>PA Linking RDs to Page</td>
<td>DA Linking RDs to Root Domain</td>
<td>Page Optimization Score</td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------------------</td>
<td>-------------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>#1</td>
<td>SEO Agency</td>
<td>Industry-Recommended Online SEO Agency</td>
<td>43</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.webfx.com/seo-agency.html">www.webfx.com/seo-agency.html</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#2</td>
<td>SEO Agency</td>
<td>Top SEO Agency</td>
<td>SEO Services - Actuate Media</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.actuatemedia.com/seo-agency/">www.actuatemedia.com/seo-agency/</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#3</td>
<td>SEO Agency</td>
<td>One Click SEO is Your Full Service SEO ...</td>
<td>35</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>oneclickseo.agency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#4</td>
<td>SEO Company</td>
<td>Search Engine Optimization Firm - SEO ...</td>
<td>61</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.seo.com">www.seo.com</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#5</td>
<td>Best SEO Companies &amp; Services - 2020 Reviews</td>
<td>Clutch.co</td>
<td>52</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>clutch.co/seo-firms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#6</td>
<td>SEO Agency</td>
<td>World-Class SEO Company</td>
<td>Directive</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>directiveconsulting.com/services/seo/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#7</td>
<td>SEO Company</td>
<td>Thrive Internet Marketing Agency</td>
<td>36</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>thriveagency.com/...rketng-services/search-engine-optimization-seo/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#8</td>
<td>SEO Agency</td>
<td>HighMark SEO Digital</td>
<td>Engineered to ...</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.highmarkseo.com">www.highmarkseo.com</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#9</td>
<td>#1 Los Angeles SEO Company</td>
<td>Search Engine Optimization</td>
<td>40</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>coalitiontechnologies.com/seo-search-engine-optimization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>Title &amp; URL</td>
<td>PA Linking RDs to Page</td>
<td>DA Linking RDs to Root Domain</td>
<td>Page Optimization Score</td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------------------</td>
<td>-------------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>#1</td>
<td>SEO Agency</td>
<td>Industry-Recommended Online SEO Agency</td>
<td>43</td>
<td>64</td>
</tr>
<tr>
<td>#2</td>
<td>SEO Agency</td>
<td>Top SEO Agency</td>
<td>SEO Services - Actuate Media</td>
<td>27</td>
</tr>
<tr>
<td>#3</td>
<td>SEO Agency</td>
<td>One Click SEO is Your Full Service SEO</td>
<td>35</td>
<td>31</td>
</tr>
<tr>
<td>#4</td>
<td>SEO Company</td>
<td>Search Engine Optimization Firm - SEO</td>
<td>61</td>
<td>52</td>
</tr>
<tr>
<td>#5</td>
<td>Best SEO Companies &amp; Services - 2020 Reviews</td>
<td>Clutch.co</td>
<td>52</td>
<td>62</td>
</tr>
<tr>
<td>#6</td>
<td>SEO Agency</td>
<td>World-Class SEO Company</td>
<td>Directive</td>
<td>37</td>
</tr>
<tr>
<td>#7</td>
<td>SEO Company</td>
<td>Thrive Internet Marketing Agency</td>
<td>36</td>
<td>43</td>
</tr>
<tr>
<td>#8</td>
<td>SEO Agency</td>
<td>HighMark SEO Digital</td>
<td>Engineered to</td>
<td>37</td>
</tr>
<tr>
<td>#9</td>
<td>#1 Los Angeles SEO Company</td>
<td>Search Engine Optimization</td>
<td>40</td>
<td>38</td>
</tr>
<tr>
<td>Rank</td>
<td>Title &amp; URL</td>
<td>PA Linking RDs to Page</td>
<td>DA Linking RDs to Root Domain</td>
<td>Page Optimization Score</td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------------------</td>
<td>-------------------------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>#1</td>
<td>SEO Agency</td>
<td>Industry-Recommended Online SEO Agency</td>
<td>43</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.webfx.com/seo-agency.html">www.webfx.com/seo-agency.html</a></td>
<td>58</td>
<td>17,871</td>
<td></td>
</tr>
<tr>
<td>#2</td>
<td>SEO Agency</td>
<td>Top SEO Agency</td>
<td>SEO Services - Actuate Media</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.actuatemedia.com/seo-agency/">www.actuatemedia.com/seo-agency/</a></td>
<td>12</td>
<td>277</td>
<td></td>
</tr>
<tr>
<td>#3</td>
<td>SEO Agency</td>
<td>One Click SEO is Your Full Service SEO</td>
<td>35</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>oneclickseo.agency</td>
<td>158</td>
<td>245</td>
<td></td>
</tr>
<tr>
<td>#4</td>
<td>SEO Company</td>
<td>Search Engine Optimization Firm - SEO</td>
<td>61</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.seo.com">www.seo.com</a></td>
<td>33,337</td>
<td>36,583</td>
<td></td>
</tr>
<tr>
<td>#5</td>
<td>Best SEO Companies &amp; Services - 2020 Reviews</td>
<td>Clutch.co</td>
<td>52</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>clutch.co/seo-firms</td>
<td>355</td>
<td>18,853</td>
<td></td>
</tr>
<tr>
<td>#6</td>
<td>SEO Agency</td>
<td>World-Class SEO Company</td>
<td>Directive</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>directiveconsulting.com/services/seo/</td>
<td>94</td>
<td>1,227</td>
<td></td>
</tr>
<tr>
<td>#7</td>
<td>SEO Company</td>
<td>Thrive Internet Marketing Agency</td>
<td>36</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>thriveagency.com/.../marketing-services/search-engine-optimization-seo/</td>
<td>175</td>
<td>2,387</td>
<td></td>
</tr>
<tr>
<td>#8</td>
<td>SEO Agency</td>
<td>HighMark SEO Digital</td>
<td>Engineered to</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td><a href="http://www.highmarkseo.com">www.highmarkseo.com</a></td>
<td>138</td>
<td>436</td>
<td></td>
</tr>
<tr>
<td>#9</td>
<td>#1 Los Angeles SEO Company</td>
<td>Search Engine Optimization</td>
<td>coalitiontechnologies.com/seo-search-engine-optimization</td>
<td>40</td>
</tr>
</tbody>
</table>
Extend the existing SERP table with new columns

```javascript
lighthouseFeatures = [
  {id: 'performance', title: 'Performance'},
  {id: 'accessibility', title: 'Accessibility'},
  {id: 'best-practices', title: 'Best Practices'},
  {id: 'seo', title: 'SEO'}
]

lighthouseFeatures.forEach(function (item) {
  $(`table.table.table-basic thead tr`).append(`<th class="table-header" role="columnheader" scope="col" style="width: 110px;"<div class="table-header-container"><span class="table-header-name">+item.title+'</span></div></th>`);
  $(`table.table.table-basic tbody tr`).append(`<td class="lighthouse '+item.id+'" colspan="1"></td>`);
});

$('td.lighthouse.performance').html(`<a class="run_lighthouse" href="#">Run Lighthouse</a>`);
```
```javascript
$(a.run_lighthouse).click(function(event) {
  targetUrl = $(this).closest('tr').find('a.external-link').attr('href');

  $(this).closest('tr').attr('data-targeturl', targetUrl);
  $.ajax({
    type: 'POST',
    crossDomain: true,
    success: function(responseData, textStatus, jqXHR) {
      requestedUrl = responseData.lhr.requestedUrl
      outputRow = $("tr[data-targeturl='"+requestedUrl+]"\n
      lighthouseFeatures.forEach(function(item) {
        outputRow.find('.id'+item.id).text(Math.round(100 * item.lhr.categories[item.id].score));
      });
    }
  });
});
```

A more customizable alternative to $.get
<table>
<thead>
<tr>
<th>Rank</th>
<th>Title &amp; URL</th>
<th>PA Linking RDs to Page</th>
<th>DA Linking RDs to Root Domain</th>
<th>Page Optimization Score</th>
<th>Performance</th>
<th>Accessibility</th>
<th>Best Practices</th>
<th>SEO</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>SEO Agency</td>
<td>Industry-Recommended Online SEO Agency - <a href="http://www.webfx.com/seo-agency.html">www.webfx.com/seo-agency.html</a></td>
<td>43</td>
<td>64</td>
<td>17,871</td>
<td>67</td>
<td>42</td>
<td>84</td>
</tr>
<tr>
<td>#2</td>
<td>SEO Agency</td>
<td>Top SEO Agency</td>
<td>SEO Services - Actuate Media - <a href="http://www.actuatemedia.com/seo-agency/">www.actuatemedia.com/seo-agency/</a></td>
<td>27</td>
<td>24</td>
<td>277</td>
<td>62</td>
<td>19</td>
</tr>
<tr>
<td>#3</td>
<td>SEO Agency</td>
<td>One Click SEO is Your Full Service SEO - oneclickseo.agency</td>
<td>35</td>
<td>31</td>
<td>245</td>
<td>78</td>
<td>15</td>
<td>86</td>
</tr>
<tr>
<td>#4</td>
<td>SEO Company</td>
<td>Search Engine Optimization Firm - SEO Company - <a href="http://www.seo.com">www.seo.com</a></td>
<td>61</td>
<td>52</td>
<td>36,583</td>
<td>79</td>
<td>12</td>
<td>66</td>
</tr>
<tr>
<td>#5</td>
<td>Best SEO Companies &amp; Services - 2020 Reviews</td>
<td>Clutch.co clutch.co/seo-firms</td>
<td>52</td>
<td>62</td>
<td>18,853</td>
<td>79</td>
<td>48</td>
<td>84</td>
</tr>
<tr>
<td>#6</td>
<td>SEO Agency</td>
<td>World-Class SEO Company</td>
<td>Directive consulting.com/services/seo/</td>
<td>37</td>
<td>45</td>
<td>1,227</td>
<td>Run Lighthouse</td>
<td></td>
</tr>
<tr>
<td>#7</td>
<td>SEO Company</td>
<td>Thrive Internet Marketing Agency</td>
<td>thriveagency.com/.../search-engine-optimization-seo/</td>
<td>36</td>
<td>43</td>
<td>2,387</td>
<td>Run Lighthouse</td>
<td></td>
</tr>
<tr>
<td>#8</td>
<td>SEO Agency</td>
<td>HighMark SEO Digital</td>
<td>Engineered to - <a href="http://www.highmarkseo.com">www.highmarkseo.com</a></td>
<td>37</td>
<td>27</td>
<td>436</td>
<td>Run Lighthouse</td>
<td></td>
</tr>
<tr>
<td>#9</td>
<td>#1 Los Angeles SEO Company</td>
<td>Search Engine Optimization</td>
<td>coalitiontechnologies.com/seo-search-engine-optimization</td>
<td>40</td>
<td>38</td>
<td>2,187</td>
<td>Run Lighthouse</td>
<td></td>
</tr>
<tr>
<td>Rank</td>
<td>Title &amp; URL</td>
<td>PA Linking RDs to Page</td>
<td>DA Linking RDs to Root Domain</td>
<td>Page Optimization Score</td>
<td>Performance</td>
<td>Accessibility</td>
<td>Best Practices</td>
<td>SEO</td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------------------------</td>
<td>------------------------</td>
<td>-------------------------------</td>
<td>-------------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>----------------</td>
<td>-----</td>
</tr>
<tr>
<td>#1</td>
<td>SEO Agency</td>
<td>Industry-Recommended Online SEO Agency ... <a href="http://www.webfx.com/seo-agency.html">www.webfx.com/seo-agency.html</a></td>
<td>43</td>
<td>64</td>
<td>17,871</td>
<td>67</td>
<td>42</td>
<td>84</td>
</tr>
<tr>
<td>#2</td>
<td>SEO Agency</td>
<td>Top SEO Agency</td>
<td>SEO Services - Actuate Media <a href="http://www.actuatemedia.com/seo-agency/">www.actuatemedia.com/seo-agency/</a></td>
<td>27</td>
<td>24</td>
<td>277</td>
<td>82</td>
<td>19</td>
</tr>
<tr>
<td>#3</td>
<td>SEO Agency</td>
<td>One Click SEO is Your Full Service SEO ... oneclickseo.agency</td>
<td>35</td>
<td>31</td>
<td>245</td>
<td>88</td>
<td>15</td>
<td>86</td>
</tr>
<tr>
<td>#4</td>
<td>SEO Company</td>
<td>Search Engine Optimization Firm - SEO ... <a href="http://www.seo.com">www.seo.com</a></td>
<td>61</td>
<td>52</td>
<td>36,583</td>
<td>79</td>
<td>12</td>
<td>66</td>
</tr>
<tr>
<td>#5</td>
<td>Best SEO Companies &amp; Services - 2020 Reviews</td>
<td>Clutch.co clutch.co/seo-firms</td>
<td>52</td>
<td>62</td>
<td>18,853</td>
<td>79</td>
<td>48</td>
<td>84</td>
</tr>
<tr>
<td>#6</td>
<td>SEO Agency</td>
<td>World-Class SEO Company</td>
<td>Directive directiveconsulting.com/services/seo/</td>
<td>37</td>
<td>45</td>
<td>1,227</td>
<td>95</td>
<td>Run Lighthouse</td>
</tr>
<tr>
<td>#7</td>
<td>SEO Company</td>
<td>Thrive Internet Marketing Agency thriveagency.com/.../marketing-services/search-engine-optimization-seo/</td>
<td>36</td>
<td>43</td>
<td>2,387</td>
<td>70</td>
<td>Run Lighthouse</td>
<td></td>
</tr>
<tr>
<td>#8</td>
<td>SEO Agency</td>
<td>HighMark SEO Digital</td>
<td>Engineered to ... <a href="http://www.highmarkseo.com">www.highmarkseo.com</a></td>
<td>37</td>
<td>27</td>
<td>436</td>
<td>96</td>
<td>Run Lighthouse</td>
</tr>
<tr>
<td>#9</td>
<td>#1 Los Angeles SEO Company</td>
<td>Search Engine Optimization coalitiontechnologies.com/seo-search-engine-optimization</td>
<td>40</td>
<td>38</td>
<td>2,187</td>
<td>81</td>
<td>Run Lighthouse</td>
<td></td>
</tr>
</tbody>
</table>
See Lighthouse metrics in your Moz Pro campaign!

Inspect the code for yourself, or install the bookmarklet from here:

ousbey.com / mozcon
Epilogue:
Your Toolkit
Run JavaScript as if it’s on the site

**Paste code into the console**

Very quick to do, especially while testing, and easy to change the code everytime. But you have to store the code somewhere to copy from, it takes a bit of work, and it’s difficult to share with other people.

**Store code in a bookmarklet**

Very easy to run code any time you want. But this can be cumbersome to update if you’re changing code frequently, and makes it harder to share with others.

**Store code in a bookmarklet**

Also very easy to run code any time you want, and to share with others. But you have to find somewhere to host your code online.
The most useful tools at your disposal

$.get() or $.post() or $.ajax()
Request content from any other URL on the site that you’re on

$('#input') or .find('h2.rating')
Use regular CSS selectors to choose an element on a page

.each(...)
Loop through every matching element you found on a page

.attr('href') or .text() or .text('new text here')
Grab content from an element on a page, or write to page
Tips for becoming an effective JavaScript hacker

Learn a little bit of JavaScript, and a lot of jQuery

Start with existing code, and edit it to understand how it works

Google everything - it’s all been solved before

Write code slowly and check / debug it a bit at a time

Ignore everything about code quality, elegance, efficiency and scalability

If it works, then it works
I'm Rob Ousbey. I spend my time thinking about customers & audiences. In my day job as VP at Distilled, I lead an incredible team of digital marketing experts, who deliver huge results by providing expert insights & advice. We support all sorts of organizations - from the world's largest companies to exciting new startups.

ousbey.com / mozcon
Thank you

ROB OUSBEY
VP, Product, Moz

@robousbey