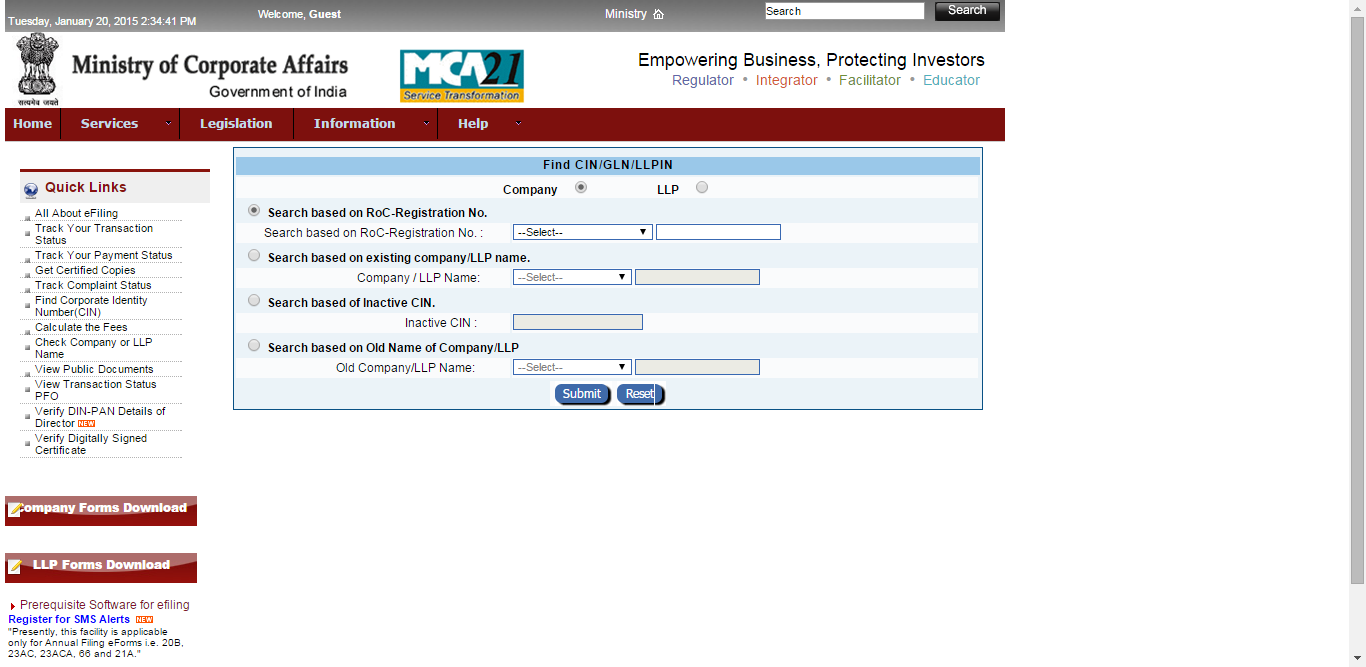
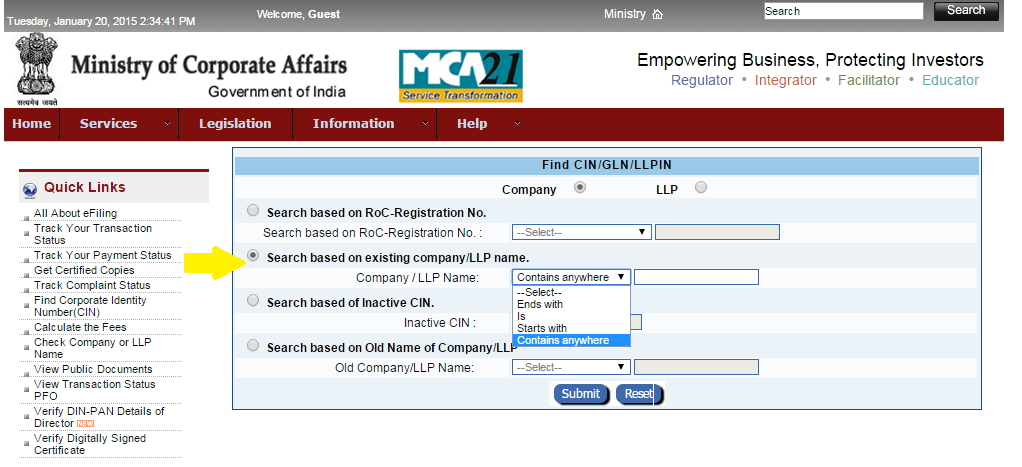
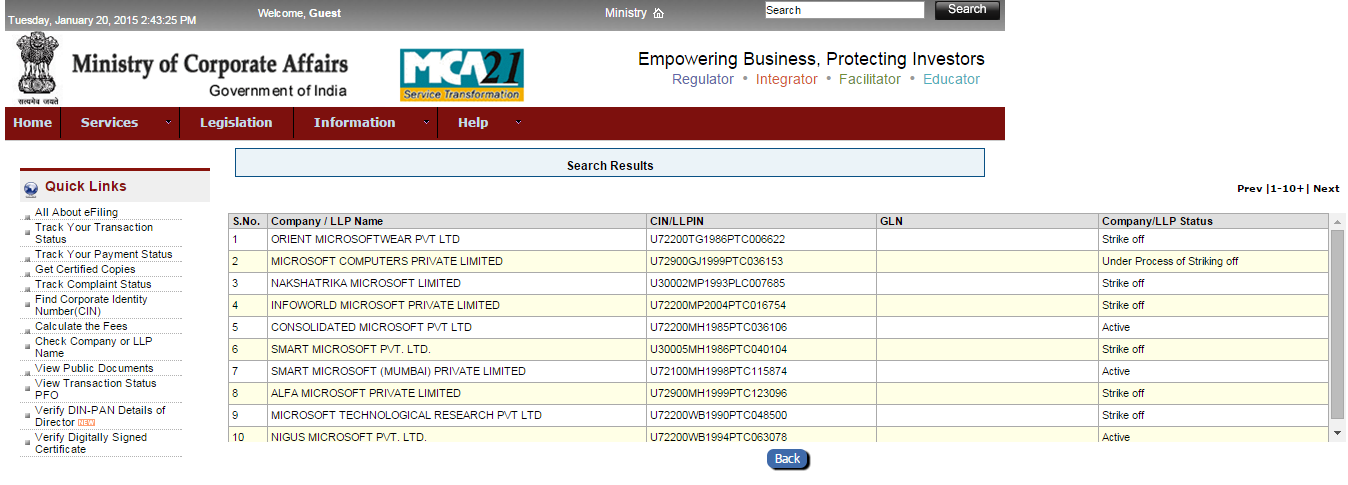
**Objective: Find the CIN/LLPIN of an Indian company given its name.**  
  
We start off [here](http://mca.gov.in/DCAPortalWeb/dca/MyMCALogin.do?method=setDefaultProperty&mode=14).

[[](http://2.bp.blogspot.com/-4swSUUsKNQ0/VL4bGtxDgCI/AAAAAAAAC5s/pbtiUn8pwrI/s1600/1.png)](http://2.bp.blogspot.com/-4swSUUsKNQ0/VL4bGtxDgCI/AAAAAAAAC5s/pbtiUn8pwrI/s1600/1.png)

Here's what we want to do step by step:  
1. Open <http://mca.gov.in/DCAPortalWeb/dca/MyMCALogin.do?method=setDefaultProperty&mode=14>  
2. Select 2nd radio button (i.e. the one just before **Search based on existing company/LLP name**)  
3. Select the **Contains anywhere** option from the dropdown menu right next to **Compnay/LLP Name.**

[](http://1.bp.blogspot.com/-waa0sUAuNkQ/VL4b509dQ4I/AAAAAAAAC54/3u-sQAF15Jk/s1600/2.png)

4. Type the name of the company you want. For now I'm not taking user input and just typing something like "Microsoft".  
5. Click the big blue Submit button.  
6. Get redirected to<http://mca.gov.in/DCAPortalWeb/dca/CompanyCINSRAction.do?method=find>

[](http://4.bp.blogspot.com/-xtzOF9mev0U/VL4c7SOyA9I/AAAAAAAAC58/bURwT8zYF4U/s1600/3.png)

7. Print out **Company/LLP name** and **CIN/LLPIN** from the table either to the console or to a file or to both.  
  
**Tools to Install:**  
  
[Python 2.6 or greater](https://www.python.org/downloads/)  
[PhantomJS 1.8.2](https://code.google.com/p/phantomjs/downloads/detail?name=phantomjs-1.8.2-windows.zip&can=4&q=)  
([PhantomJS 1.9](http://phantomjs.org/" \t "_blank) is out but I downgraded since it gives me errors with CasperJS 1.1 and above. You could try downloading the latest version though).  
[CasperJS](http://casperjs.readthedocs.org/en/latest/installation.html)  
**Optional:** If you have Chrome, [Selector Gadget](https://chrome.google.com/webstore/detail/selectorgadget/mhjhnkcfbdhnjickkkdbjoemdmbfginb?hl=en). If you have Firefox, [Firebug](https://addons.mozilla.org/en-US/firefox/addon/firebug/) and [Firefinder](https://addons.mozilla.org/en-us/firefox/addon/firefinder-for-firebug/). These are solely to help you find out element names and ids on the web page, you can find them through Inspect Element (right click on any webpage, scroll down to **Inspect Element**) on Chrome or Firefox as well.   
  
**Step by Step**  
First we create the casper object. I've set verbose to true so you get very detailed debugging on the console to help you understand what's happening step by step.

var casper = require('casper').create({

verbose: true,

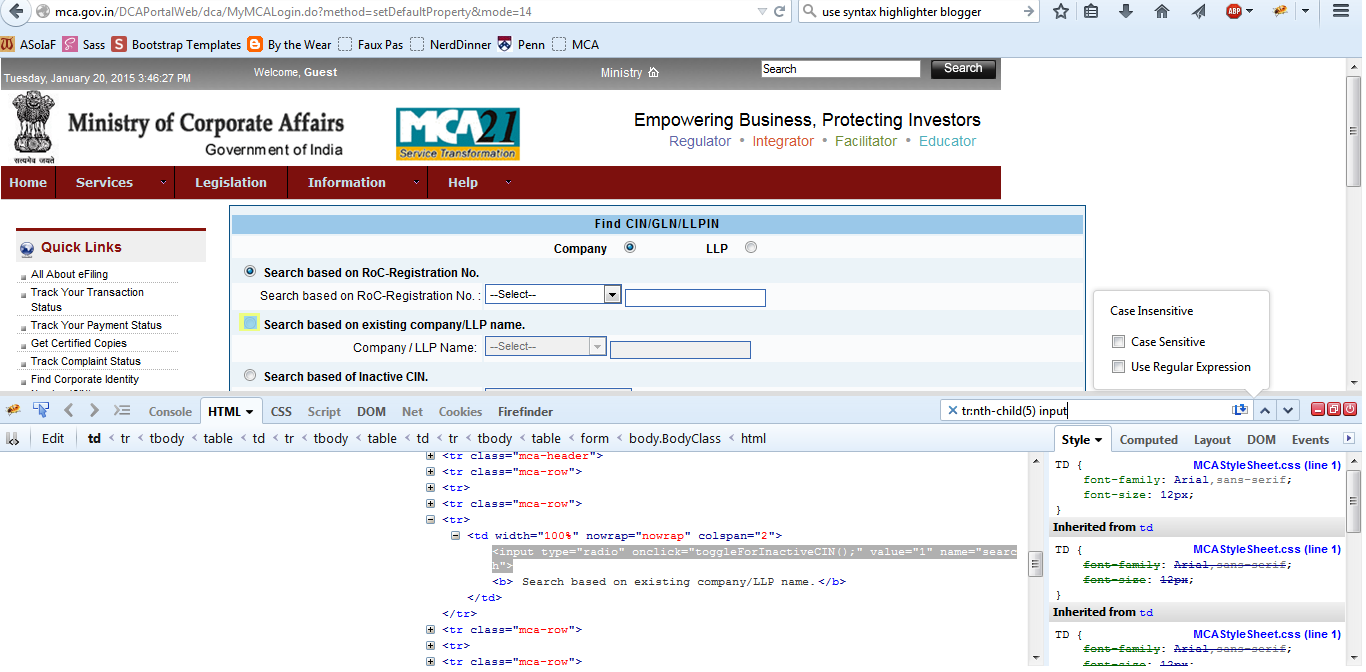
logLevel: "debug"

});

We open the desired page.

casper.start('http://mca.gov.in/DCAPortalWeb/dca/MyMCALogin.do?method=setDefaultProperty&mode=14');

The next step is a bit tricky. We want to find the id of the 2nd radio button on the page so we can [click](http://casperjs.readthedocs.org/en/latest/modules/casper.html#click) on it. I've used Firebug in Firefox for this. Here's a screenshot:

[](http://3.bp.blogspot.com/-_KtW_jHWUnc/VL4r4Urwz-I/AAAAAAAAC6M/ddiRTTRmUHo/s1600/4.png)

Expect to have to fiddle around a bit and search for relevant StackOverflow questions to help you find the id of your desired DOM element. Ultimately, I found out that the id of the element I want is: **tr:nth-child(5) input**  
  
So now we click on it:

casper.then(function() {

this.click('tr:nth-child(5) input');

});

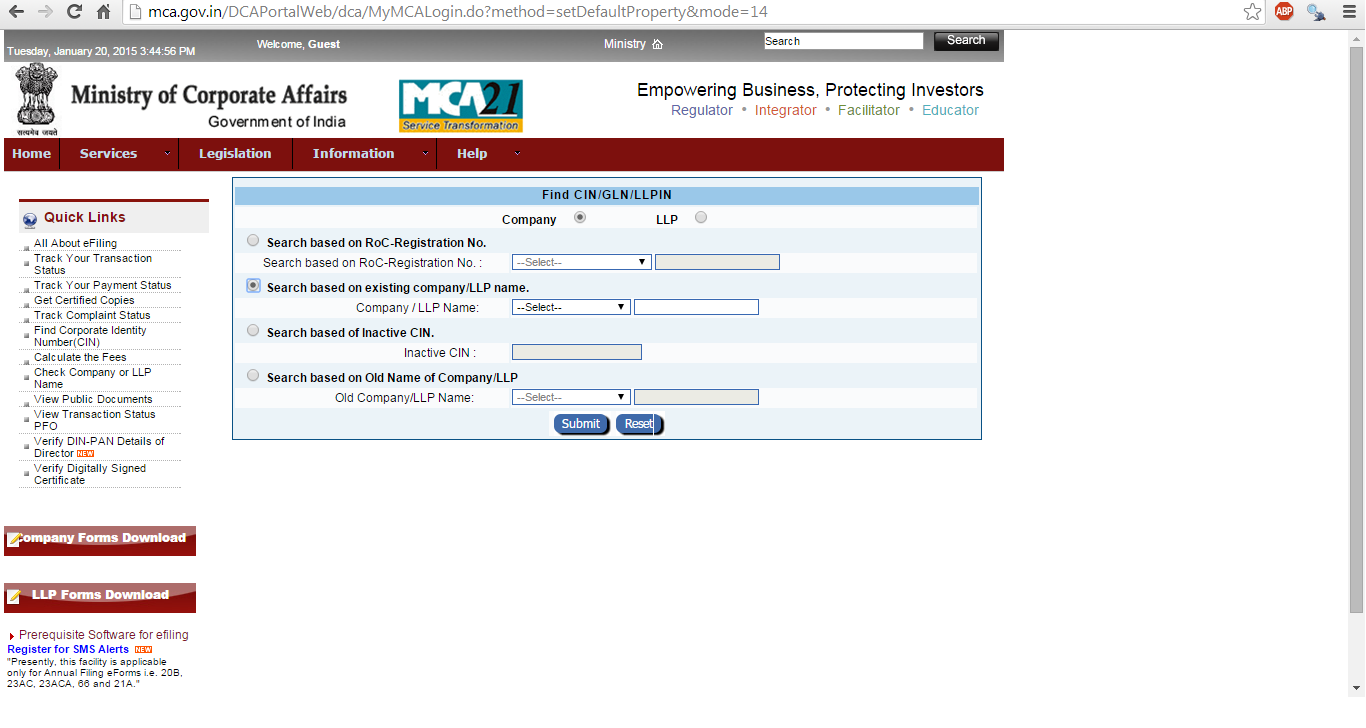
At this point you might want to take a screenshot with this command:

casper.then(function() {

this.capture("result1.png");

});

That's one of the best features of CasperJS, you can take a screenshot whenever you want to see what's happening. You should get this sort of image:

[](http://2.bp.blogspot.com/-CYYRMf8QKkQ/VL4s1lqqMrI/AAAAAAAAC6U/l-HV0fxgENw/s1600/5.png)

Next we want to find out the name of the drop-down menu right next to **Company/LLP Name.**  
Opening the page source gives us this handy code snippet, waaaay down:

<td class="Field" width="63%" nowrap="">  
<select id="searchCriteria" class="cb" size="" name="searchCriteria">  
<option value="">--Select--</option>  
<option value="ENDS">Ends with</option>  
<option value="IS">Is</option>  
<option value="STRT">Starts with</option>  
<option value="CONT">Contains anywhere</option>  
</select>  
<input id="companyName" class="txt" type="text" onblur="alphaNumericFieldCheck(document.forms[0].companyName,'NSCWCNMS');isValidCompDtls();checkMinLength(document.forms[0].companyName,2);" value="" maxlength="100" name="companyName">  
</td>

Thus, the id and name of our dropdown menu is **searchCriteria**. And the id of the box right next to it, which we want to fill up with a name, is **companyName.** To click on the dropdown menu:

casper.then(function() {

this.click('#searchCriteria');

});

The option we want to select has the id CONT - so we can search for companies with any part of a name that we can think of. We also want to fill out the box next to it and we can do both with CasperJS's [fill() method](http://casperjs.readthedocs.org/en/latest/modules/casper.html#fill).

casper.then(function() {

this.fill('form[name="CompanyCINSRForm"]', {

'searchCriteria' : 'CONT',

'companyName' : 'Microsoft'

}, false);

});

We set Boolean submit to false since the page causes small errors and does not redirect correctly when we set it to true.   
Next we want to click on the Submit button at the bottom of the page, which has id **#Default.**

casper.then(function() {

this.click('#Default');

});

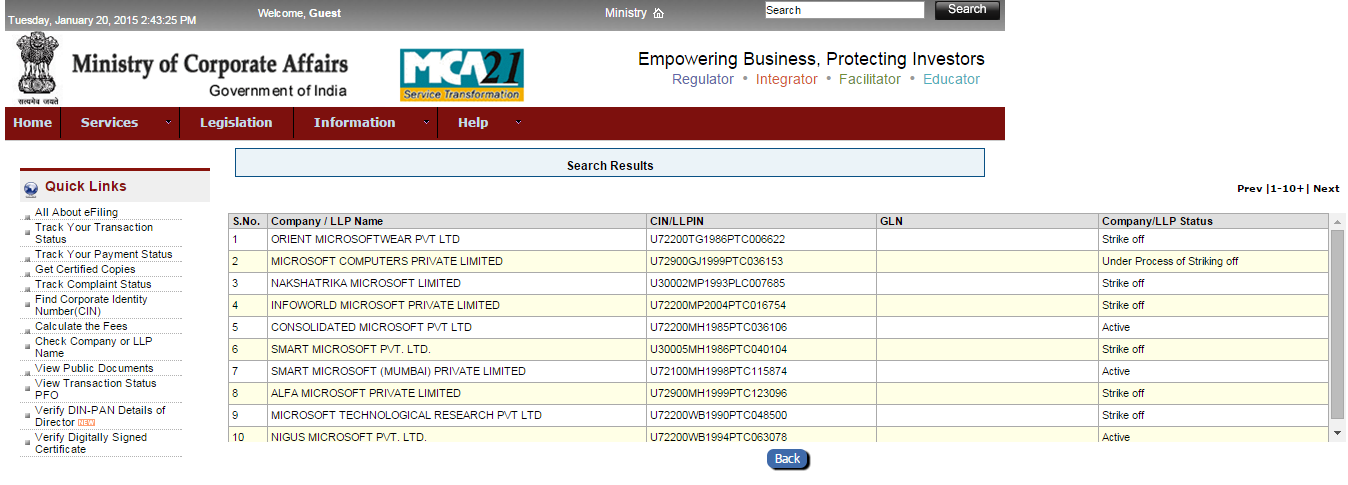
Here's a helper function to tell you where you're redirected to after clicking the Submit button (I've commented it out in the main code).

casper.then(function() {

console.log('clicked Submit, new location is ' + this.getCurrentUrl());

});

Now we should be here:

[](http://4.bp.blogspot.com/-xtzOF9mev0U/VL4c7SOyA9I/AAAAAAAAC6A/albf8-iqq58/s1600/3.png)

We want to print out Column 2 and Column 3. What I've basically done is replicated Victor Yee's code (which he used to print the names of towns in Vermont, again using CasperJS) with a few modifications. You can see his very clearly-explained code on his blog [here](http://www.victorwyee.com/js/webscraping-with-casperjs-phantomjs/).  
  
My code is:

casper.then(function() {

this.echo(this.getTitle());

// Get info on all elements matching this CSS selector

var name\_selector = 'table[id="list1"] tbody tr td:nth-of-type(2)';

var cin\_selector = 'table[id="list1"] tbody tr td:nth-of-type(3)';

var names\_info = this.getElementsInfo(name\_selector); // an array of object literals

var cin\_info = this.getElementsInfo(cin\_selector);

var names = [];

var cin = [];

for (var i = 0; i < names\_info.length; i++) {

names.push(names\_info[i].text);

cin.push(cin\_info[i].text);

}

require('utils').dump(names);

require('utils').dump(cin);

var fs = require('fs');

var myfile = "output.txt"

fs.write(myfile, cin, 'w');

});

I've printed out the company names and CINs to the console and also printed out CINs (not formatted at all, as you can see) to 'output.txt'.