Web Scraping in Python using Selenium

# Problem statement

*The client wants to get contact number from name. So he is wanted to use the MTNL/ BSNL/ LinkedIN site(s) for parsing.*

*BSNL:*

*Seems to be easier to crack. Search by number is working but search by Name is running into issues. Try with Python Selenium or RoboBrowser to get good parsing results.*

*http://www.bsnl.in/opencms/bsnl/BSNL/directory\_services/online\_telephone\_directory.html*

*from that Maharashtra site is : http://www.bsnl.in/opencms/bsnl/BSNL/directory\_services/online\_telephone\_directory.html*

*use last name search. Test files are attached as reference but you can choose your own method.*

*MTNL:*

*http://mtnlmumbai.in/index.php/directory-search*

*Has a captcha to solve. Used a brute method to solve but still has problems. (we can view this later).*

*LinkedIN: has api but not able to get phone number or email address. Try brute scraping and see if that works for you.*

# Implementation

## Installation

* Install selenium (do not give any of your files name as ‘selenium’, else web driver error)
* Download and keep geckodriver.exe in PATH
* Run Sample test

import time

from selenium import webdriver

driver = webdriver.Firefox()

time.sleep(5)

driver.quit()