Name: Saket Choudhary
USCID: 2170058637
Website Indexed: Yahoo

I used Python for this programming assignment throughout for all the the three tasks: a) Spelling Correction b) Autocomplete c) Snippet generation

Spelling Correction

- I use Peter Norvig's original implementation: 'spell.py'
- For creating `big.txt` required by 'spell.py', I parsed the text content of all html files for Yahoo using beautifulsoup and checked if they were potentially valid words (that is they contained only characters a-z|A-Z); this file is 'create_yahoo_big.py' in my submission
- I used one function from 'spell.py': 'correct()' to generate a spell-corrected word for an input word

Autocomplete

- I used flask for implementing the webservice alongside jquery's 'autocomplete()' function
- For my flask app, I created a method 'autocomplete' that would obtain query words using a HTTP GET request and respond a JSON object with list of suggestions as obtained from SOLR's '/suggest' interface; the underlying method uses a Python wrapper called solrpy
- At the jquery end, I split the query by a space and only use the last term (which will be the only term if there are no spaces) for doing an AJAX call to the '/autocomplete' API endpoint. To allow a cross-domain AJAX request I forced the return datatype to be 'jsonp' and set the jsonp format to 'json.wrf" to handle SOLR's JSON response.
- The suggestions returned by '/autocomplete' endpoint are then prefixed with rest of the preceding query terms and displayed in the dropdown
- On pressing 'submit' on the query page, the response is sent to another endpoint '/' that sends back the query results as html. If the last query term can be spell-corrected, it additionally sends another html object with "Did you mean XYZ?" where XYZ is the corrected spelling of given query

Snippet

- Once results are returned by SOLR for any query, I use the docid to parse the entire html and extract the text only section
- I split the text by a "." and call it a sentence
- I then lookup by a regular expression match for the query inside each of the sentence and return the sentence with most number of matching terms with the query
- The matching query terms are boldfaced in the snippet

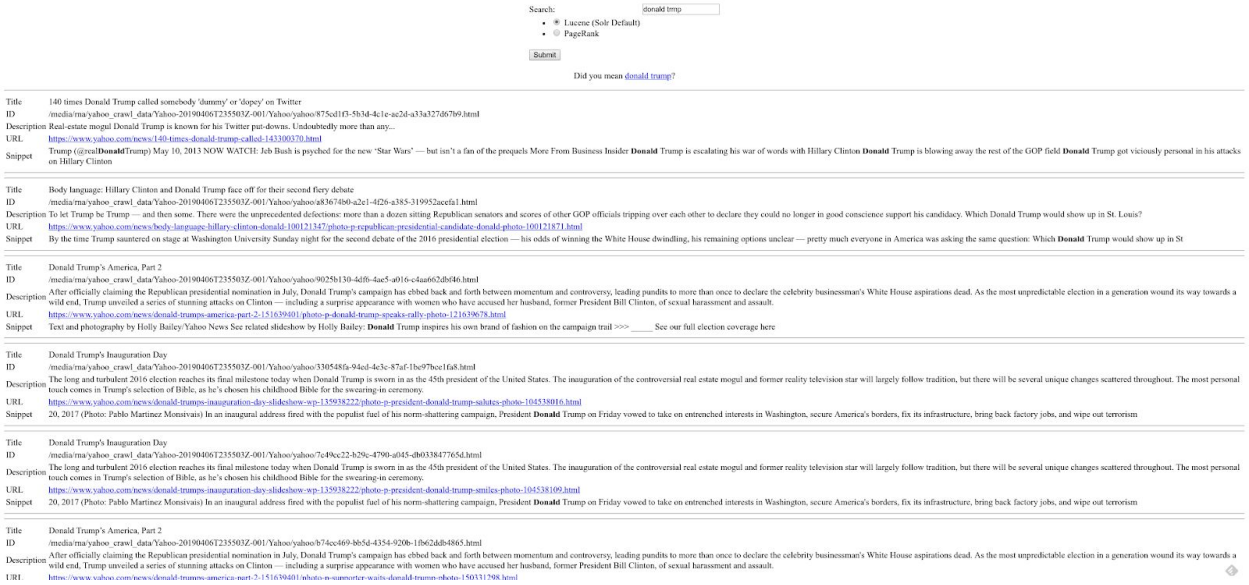
Code:

- 'templates/search.html': Tempalte for the html page for search
- 'app.py': FLASK app for all webservice
- 'spell.py': Peter Norvig's spell.py
- 'create_yahoo_big.py': Create yahoo_big.txt required for spell.py

Analysis of the results

Examples of spell-correction

1. Narendra mofi => narendra modi



Search: donald tmp

- Lucene (Solr Default)
- PageRank

Submit

Did you mean donald trump?

Title	140 times Donald Trump called somebody 'dummy' or 'dopey' on Twitter
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/875cd1f3-5b3d-4c1e-ae2d-a33a327d67b9.html
Description	Real-estate mogul Donald Trump is known for his Twitter put-downs. Undoubtedly more than any...
URL	https://www.yahoo.com/news/140-times-donald-trump-called-143300370.html
Snippet	Trump (@realDonaldTrump) May 10, 2013 NOW WATCH: Jeb Bush is psyched for the new 'Star Wars' — but isn't a fan of the prequels More From Business Insider **Donald** Trump is escalating his war of words with Hillary Clinton **Donald** Trump is blowing away the rest of the GOP field **Donald** Trump got viciously personal in his attacks on Hillary Clinton

Title	Body language: Hillary Clinton and Donald Trump face off for their second fiery debate
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/a83674b0-a2e1-4f26-a385-319952aeefa1.html
Description	To let Trump be Trump — and then some. There were the unprecedented defections: more than a dozen sitting Republican senators and scores of other GOP officials tripping over each other to declare they could no longer in good conscience support his candidacy. Which Donald Trump would show up in St. Louis?
URL	https://www.yahoo.com/news/body-language-hillary-clinton-donald-100121347/photo-p-republican-presidential-candidate-donald-photo-100121871.html
Snippet	By the time Trump sauntered on stage at Washington University Sunday night for the second debate of the 2016 presidential election — his odds of winning the White House dwindling, his remaining options unclear — pretty much everyone in America was asking the same question: Which **Donald** Trump would show up in St

Title	Donald Trump's America, Part 2
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/9025b130-4df6-4ae5-a016-c4aa662dbf46.html
Description	After officially claiming the Republican presidential nomination in July, Donald Trump's campaign has ebbed back and forth between momentum and controversy, leading pundits to more than once to declare the celebrity businessman's White House aspirations dead. As the most unpredictable election in a generation wound its way towards a wild end, Trump unveiled a series of stunning attacks on Clinton — including a surprise appearance with women who have accused her husband, former President Bill Clinton, of sexual harassment and assault.
URL	https://www.yahoo.com/news/donald-trumps-america-part-2-151639401/photo-p-donald-trump-speaks-rally-photo-121639678.html
Snippet	Text and photography by Holly Bailey/Yahoo News See related slideshow by Holly Bailey: **Donald** Trump inspires his own brand of fashion on the campaign trail >>> _____ See our full election coverage here

Title	Donald Trump's Inauguration Day
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/330548fa-94cd-4e3c-87af-1be97bce1fa8.html
Description	The long and turbulent 2016 election reaches its final milestone today when Donald Trump is sworn in as the 45th president of the United States. The inauguration of the controversial real estate mogul and former reality television star will largely follow tradition, but there will be several unique changes scattered throughout. The most personal touch comes in Trump's selection of Bible, as he's chosen his childhood Bible for the swearing-in ceremony.
URL	https://www.yahoo.com/news/donald-trumps-inauguration-day-slideshow-wp-135938222/photo-p-president-donald-trump-salutes-photo-104538016.html
Snippet	20, 2017 (Photo: Pablo Martinez Monsivais) In an inaugural address fired with the populist fuel of his norm-shattering campaign, President **Donald** Trump on Friday vowed to take on entrenched interests in Washington, secure America's borders, fix its infrastructure, bring back factory jobs, and wipe out terrorism

Title	Donald Trump's Inauguration Day
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/7c49ce22-b29c-4790-a045-db033847765d.html
Description	The long and turbulent 2016 election reaches its final milestone today when Donald Trump is sworn in as the 45th president of the United States. The inauguration of the controversial real estate mogul and former reality television star will largely follow tradition, but there will be several unique changes scattered throughout. The most personal touch comes in Trump's selection of Bible, as he's chosen his childhood Bible for the swearing-in ceremony.
URL	https://www.yahoo.com/news/donald-trumps-inauguration-day-slideshow-wp-135938222/photo-p-president-donald-trump-smiles-photo-104538109.html
Snippet	20, 2017 (Photo: Pablo Martinez Monsivais) In an inaugural address fired with the populist fuel of his norm-shattering campaign, President **Donald** Trump on Friday vowed to take on entrenched interests in Washington, secure America's borders, fix its infrastructure, bring back factory jobs, and wipe out terrorism

Title	Donald Trump's America, Part 2
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/b74ce469-bb5d-4354-920b-1fb62ddb4865.html
Description	After officially claiming the Republican presidential nomination in July, Donald Trump's campaign has ebbed back and forth between momentum and controversy, leading pundits to more than once to declare the celebrity businessman's White House aspirations dead. As the most unpredictable election in a generation wound its way towards a wild end, Trump unveiled a series of stunning attacks on Clinton — including a surprise appearance with women who have accused her husband, former President Bill Clinton, of sexual harassment and assault.
URL	https://www.yahoo.com/news/donald-trumps-america-part-2-151639401/photo-p-supporter-waits-donald-trump-photo-150331298.html

2. rahul gandhei => rahul gandhi

Title	U.S. Congress
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/aa7cc3c7-63c2-4dde-9551-180640290fdc.html
Description	
URL	https://www.yahoo.com/news/topics/us-congress
Snippet	"The Abbas family's story isn't new – it's Business Standard India If voted to power Congress will fill all govt posts lying vacant: **Rahul** Lamenting that Odisha has become an "unemployment centre" under the BJD government, Congress chief **Rahul** Gandhi Friday promised to fill the 1

Title	U.S. government
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/858c294b-3670-479c-903f-5d68d8a024c9.html
Description	
URL	https://www.yahoo.com/news/topics/elections-aftermath
Snippet	Calling **Rahul** Gandhi the "next Prime Minister of India", the Congress legislature party leader recalled his grandmother Indira Gandhi and mother Sonia Gandhi contesting the Lok Sabha polls from Karnataka in the past

3. Donald trmp => donal trump

Title	140 times Donald Trump called somebody 'dummy' or 'dopey' on Twitter
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/875cd1f3-5b3d-4c1e-ae2d-a33a327d7b9.html
Description	Real-estate mogul Donald Trump is known for his Twitter put-downs. Undoubtedly more than any...
URL	https://www.yahoo.com/news/140-times-donald-trump-called-143300370.html
Snippet	Trump (@realDonaldTrump) May 10, 2013 NOW WATCH: Jeb Bush is psyched for the new 'Star Wars' — but isn't a fan of the prequels More From Business Insider **Donald** Trump is escalating his war of words with Hillary Clinton **Donald** Trump is blowing away the rest of the GOP field **Donald** Trump got viciously personal in his attacks on Hillary Clinton

Title	Body language: Hillary Clinton and Donald Trump face off for their second fiery debate
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/a83674b0-a2e1-4f26-a385-319952aeefa1.html
Description	To let Trump be Trump — and then some. There were the unprecedented defections: more than a dozen sitting Republican senators and scores of other GOP officials tripping over each other to declare they could no longer in good conscience support his candidacy. Which Donald Trump would show up in St. Louis?
URL	https://www.yahoo.com/news/body-language-hillary-clinton-donald-100121347/photo-p-republican-presidential-candidate-donald-photo-100121871.html
Snippet	By the time Trump sauntered on stage at Washington University Sunday night for the second debate of the 2016 presidential election — his odds of winning the White House dwindling, his remaining options unclear — pretty much everyone in America was asking the same question: Which **Donald** Trump would show up in St

Title	Donald Trump's America, Part 2
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/9025b130-4df6-4ae5-a016-c4aa662dbf46.html
Description	After officially claiming the Republican presidential nomination in July, Donald Trump's campaign has ebbed back and forth between momentum and controversy, leading pundits to more than once to declare the celebrity businessman's White House aspirations dead. As the most unpredictable election in a generation wound its way towards a wild end, Trump unveiled a series of stunning attacks on Clinton — including a surprise appearance with women who have accused her husband, former President Bill Clinton, of sexual harassment and assault.
URL	https://www.yahoo.com/news/donald-trumps-america-part-2-151639401/photo-p-donald-trump-speaks-rally-photo-121639678.html
Snippet	Text and photography by Holly Bailey/Yahoo News See related slideshow by Holly Bailey: **Donald** Trump inspires his own brand of fashion on the campaign trail >>> _____ See our full election coverage here

Title	Donald Trump's Inauguration Day
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/330548fa-94ed-4e3c-87af-1be97bee1fa8.html
Description	The long and turbulent 2016 election reaches its final milestone today when Donald Trump is sworn in as the 45th president of the United States. The inauguration of the controversial real estate mogul and former reality television star will largely follow tradition, but there will be several unique changes scattered throughout. The most personal touch comes in Trump's selection of Bible, as he's chosen his childhood Bible for the swearing-in ceremony.
URL	https://www.yahoo.com/news/donald-trumps-inauguration-day-slideshow-wp-135938222/photo-p-president-donald-trump-salutes-photo-104538016.html
Snippet	20, 2017 (Photo: Pablo Martinez Monsivais) In an inaugural address fired with the populist fuel of his norm-shattering campaign, President **Donald** Trump on Friday vowed to take on entrenched interests in Washington, secure America's borders, fix its infrastructure, bring back factory jobs, and wipe out terrorism

Title	Donald Trump's Inauguration Day
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/7c49ce22-b29c-4790-a045-db033847765d.html
Description	The long and turbulent 2016 election reaches its final milestone today when Donald Trump is sworn in as the 45th president of the United States. The inauguration of the controversial real estate mogul and former reality television star will largely follow tradition, but there will be several unique changes scattered throughout. The most personal touch comes in Trump's selection of Bible, as he's chosen his childhood Bible for the swearing-in ceremony.
URL	https://www.yahoo.com/news/donald-trumps-inauguration-day-slideshow-wp-135938222/photo-p-president-donald-trump-smiles-photo-104538109.html
Snippet	20, 2017 (Photo: Pablo Martinez Monsivais) In an inaugural address fired with the populist fuel of his norm-shattering campaign, President **Donald** Trump on Friday vowed to take on entrenched interests in Washington, secure America's borders, fix its infrastructure, bring back factory jobs, and wipe out terrorism

Title	Donald Trump's America, Part 2
ID	/media/rna/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/b74ec469-bb5d-4354-920b-1fb62ddb4865.html
Description	After officially claiming the Republican presidential nomination in July, Donald Trump's campaign has ebbed back and forth between momentum and controversy, leading pundits to more than once to declare the celebrity businessman's White House aspirations dead. As the most unpredictable election in a generation wound its way towards a wild end, Trump unveiled a series of stunning attacks on Clinton — including a surprise appearance with women who have accused her husband, former President Bill Clinton, of sexual harassment and assault.
URL	https://www.yahoo.com/news/donald-trumps-america-part-2-151639401/photo-p-supporter-waits-donald-trump-photo-150331298.html

4. **Dimey => dismay**

5. **philosphy => philosophy**

Clicking philosophy results in the following set of results:

Title	What does Neil Gorsuch's book on suicide reveal about his philosophy?
ID	/media/rma/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/b1e2847e-3d4e-4924-b829-42f0272deb14.html
Description	The "options" it wants to expand are for terminally ill patients to obtain a lethal dose of barbiturates to take if and when they decide to die. This is an issue on which Gorsuch is an expert, as the author of a book on the subject, an essay on constitutional law and moral philosophy titled "The Future
URL	https://www.yahoo.com/news/neil-gorsuch-on-the-future-of-assisted-suicide-100028261.html
Snippet	com-Chip Somodevilla/Getty Images) More The crux of Gorsuch's argument, which he pursues through thousands of years of Western **philosophy** and centuries of American jurisprudence, is summed up pithily in his first chapter: the idea "that all human beings are intrinsically valuable and the intentional taking of life by private persons is always wrong

Title	Conservative intellectuals on Donald Trump's philosophy
ID	/media/rma/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/6d9befd8-6585-414e-a3fb-bbe59d3f0910.html
Description	President Trump's unbridled populism has fanned the flames of resentment against intellectuals, including conservatives. Yahoo News reached out to several prominent conservative thinkers to see what they make of the Trump spectacle and what it portends about the future of conservatism.
URL	https://www.yahoo.com/news/cying-heart-conservatism-trump-era-090023043.html
Snippet	But he sees a divide here: mainstream American conservatives talk like Randians (Russian-American novelist Ayn Rand espoused a **philosophy** of individualism and radical self-interest) but live like Tocquevillians (French political theorist Alexis de Tocqueville greatly admired American democracy but worried about a "tyranny of the majority" trampling individual rights)

Title	Democrats try to co-opt populist rage. Hilarity ensues.
ID	/media/rma/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/082e8143-9201-41d5-94f1-66b27385f2e4.html
Description	The problem with the Democrats' new agenda isn't its language or gauziness. It's the underlying philosophy, which misreads the Trump campaign's core appeal.
URL	https://www.yahoo.com/news/democrats-try-co-opt-populist-rage-hilarity-ensues-050005021.html
Snippet	It's more about the underlying **philosophy**, which misreads in some fundamental way the core appeal of Trump's campaign

Title	Gorsuch frustrates Democrats at confirmation hearing [Video]
ID	/media/rma/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/a8a5ec72-c220-4bea-92d1-1e5434774b1d.html
Description	During a 10-hour grilling from senators Tuesday, Judge Neil Gorsuch offered few hints as to his judicial philosophy, frustrating the Judiciary Committee's Democrats in a polished and calm performance. Gorsuch — sprinkling his answers to the committee's questions with "gosh" and "golly" and "goodness
URL	https://www.yahoo.com/news/gorsuch-frustrates-democrats-at-confirmation-hearing-015048427.html
Snippet	Home Mail News Finance Sports Entertainment Search Mobile More Yahoo Search Search News Search web Skip to Navigation Skip to Main Content Skip to Related Content Mail Gorsuch frustrates Democrats at confirmation hearing Liz Goodwin Senior National Affairs Reporter , Yahoo News • March 22, 2017 WASHINGTON — During a 10-hour grilling from senators Tuesday, Judge Neil Gorsuch offered few hints as to his judicial **philosophy**, frustrating the Judiciary Committee's Democrats in a polished and calm performance

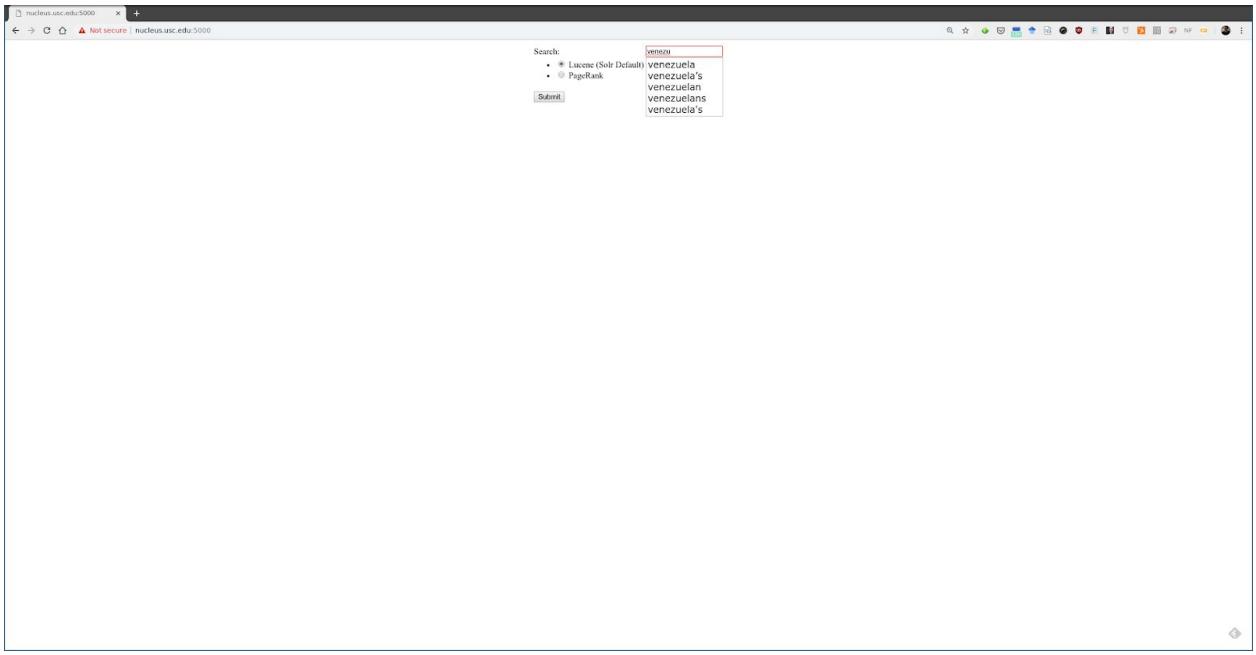
Title	Trump's so-what climate change policy
ID	/media/rma/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/3b14955b-6362-4f12-80e2-c26b7798aee7.html
Description	Yahoo's Jerry Adler examines the Trump administration's so-what climate change policy.
URL	https://www.yahoo.com/news/trumps-future-damned-climate-change-policy-090015859.html
Snippet	Someone with the " very good brain " he likes to boast about — his uncle was a professor at MIT , in case you didn't know — undoubtedly is well versed in the **philosophy** of Nietzsche

Title	How Brett Kavanaugh's Catholic childhood shaped him
ID	/media/rma/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/d3f22e6d-18ba-49dc-bb17-94f0bd54d915.html
Description	Brett Kavanaugh's Catholic upbringing and education at Mater Dei elementary school for boys and Georgetown Prep in the suburbs of Washington, D.C., had a tremendous impact on the man he would become.
URL	https://www.yahoo.com/news/brett-kavanaughs-religious-upbringing-shaped-thinking-131552610.html
Snippet	Applying his Jesuit training Kavanaugh's judicial **philosophy** is also clearly informed by his faith

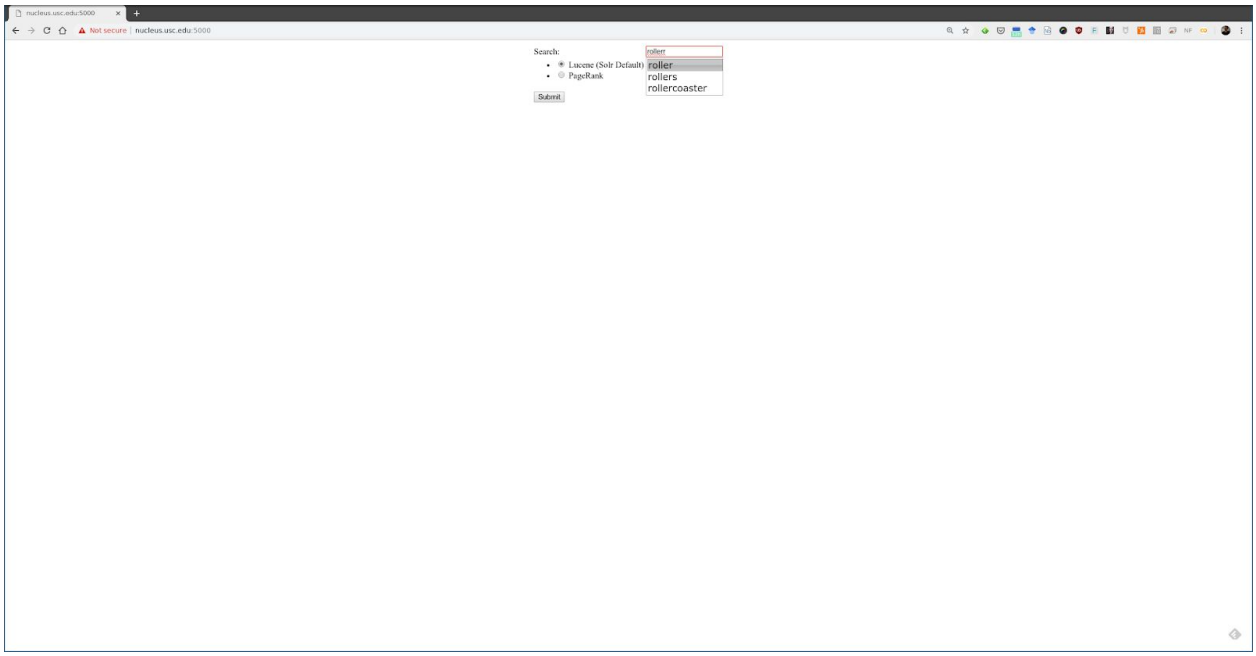
Title	New Kavanaugh allegations: It's all about politics
ID	/media/rma/yahoo_crawl_data/Yahoo-20190406T235503Z-001/Yahoo/yahoo/65d11617-6a7a-4e72-addc-27d3ec388e26.html
Description	The accusations of sexual assault against Supreme Court nominee Brett Kavanaugh raise questions about the responsibility of mature adults for events in the distant past. But the answers will be provided by politics, not moral philosophy.

Autosuggest

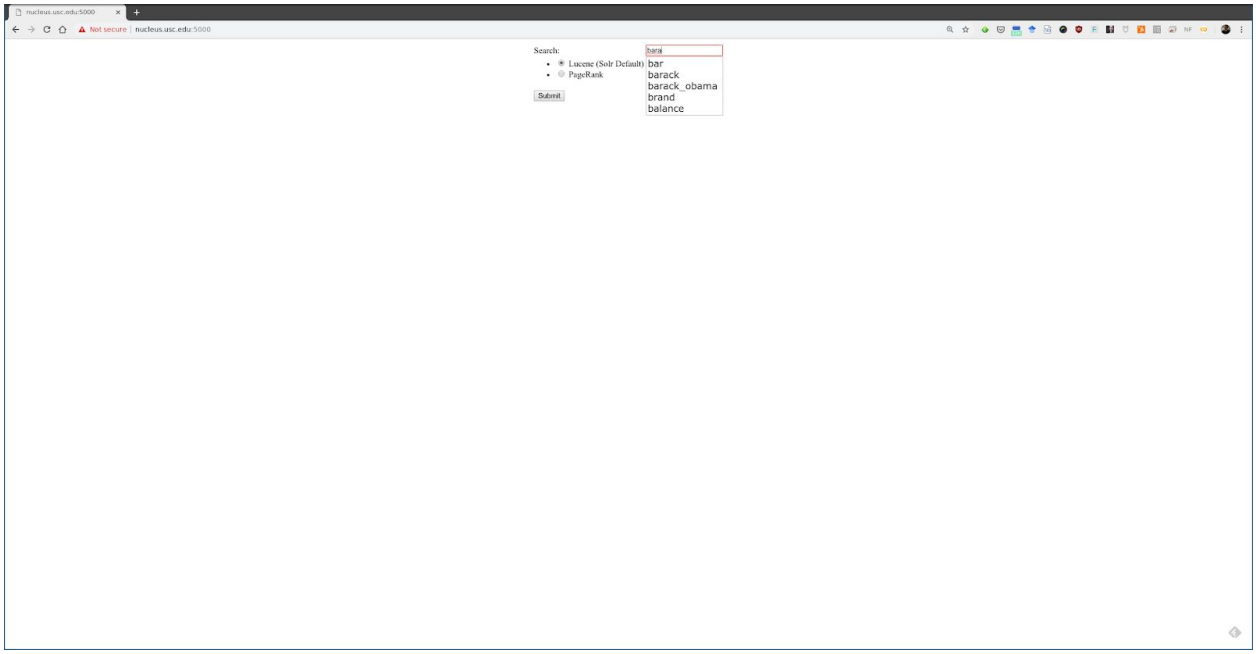
1. Venezuela



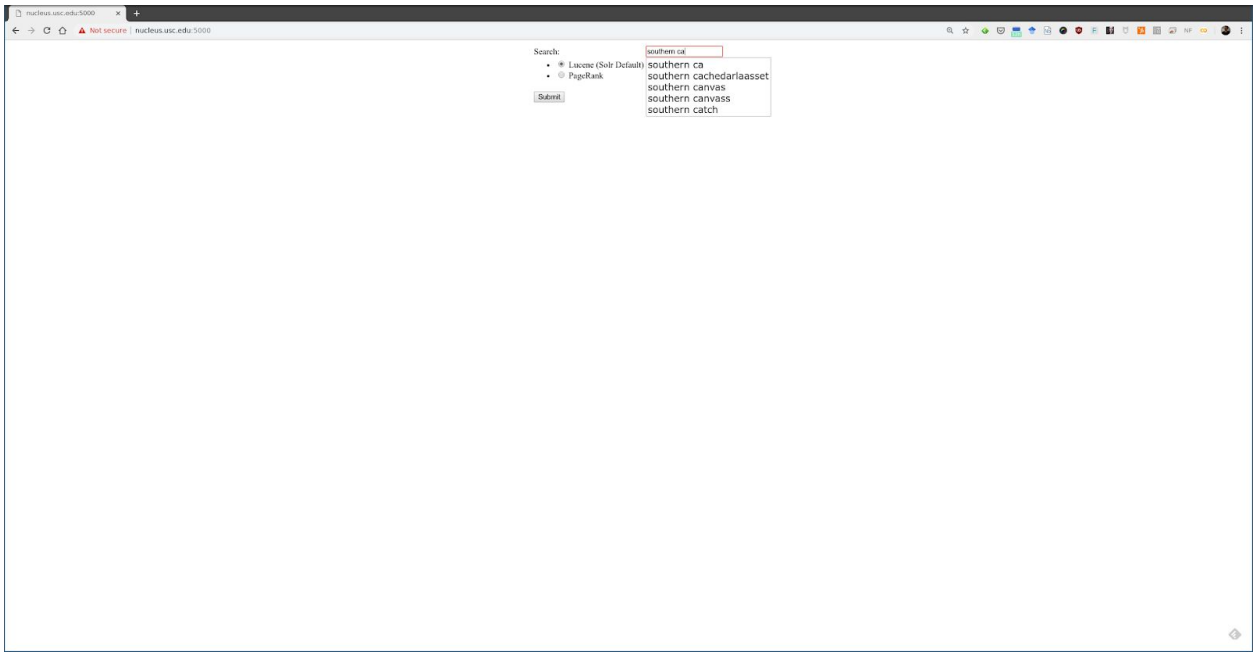
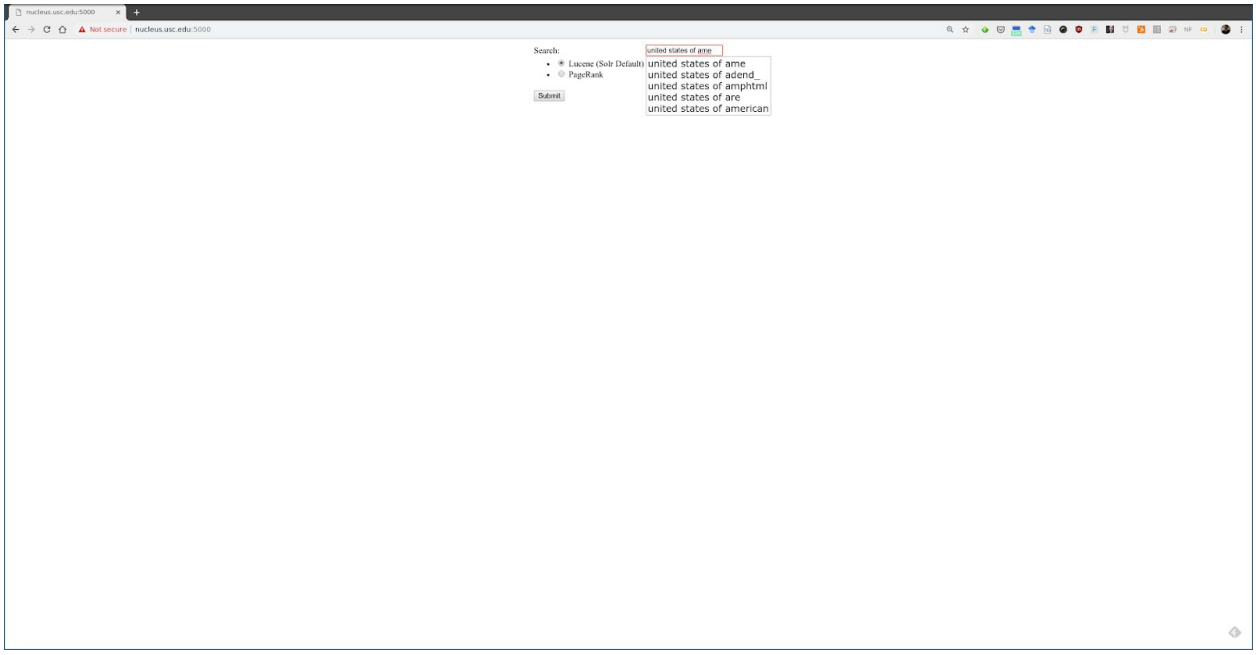
2. Rollercoaser



3. Barack Obama



4. Southern california



5. United States of America