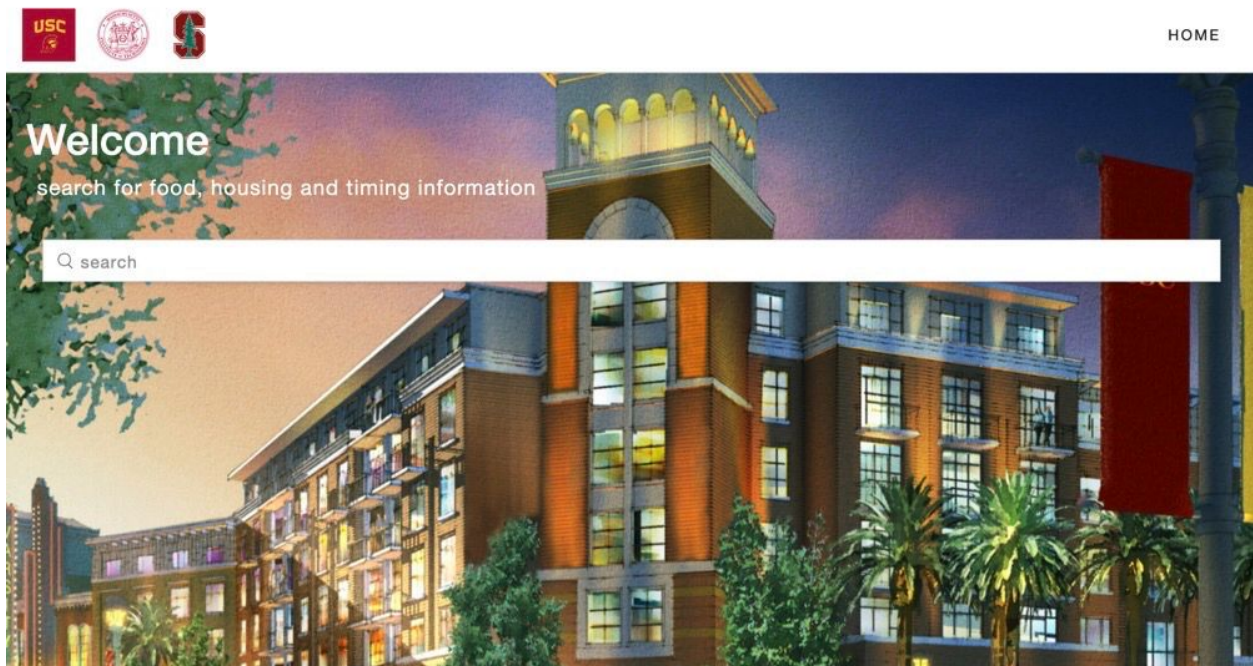# INF558 REPORT

Knowledge Graph -- A website for fresh students

Ruotian Jiang | 8389636738

Abhishek Bhatt | 6737460912

04.26.2019

INTRODUCTION

We have built a knowledge graph to assist new students to transitioning to college life as well as to make existing students life more easier by providing a unified source of all commonly required information. Information about housing, calendar and dining services from USC, MIT and stanford are provided at a single location by our application.

Data Crawling

In an effort to get as much related information as we can we had crawled data from following domains.

University	Domain	Link
USC	Calendar	https://academics.usc.edu/calendar/academic-calendar-2018-2019/
USC	Housing	https://housing.usc.edu/
USC	Dinning	https://hospitality.usc.edu/residential-dining/
MIT	Calendar	https://registrar.mit.edu/calendar/current-key-dates
MIT	Housing	https://studentlife.mit.edu/housing
MIT	Dinning	
Stanford	Calendar	https://registrar.stanford.edu/academic-calendar
Stanford	Dinning	https://rde.stanford.edu/dining

Stanford	Housing	https://rde.stanford.edu/studenthousing

Wrapper

We have a variety of wrapper to extract the information from the crawled pages . The output of the wrapper is stored in json format.

1. Calendar

```
[
   {
      "date": "....",
      "day": "Monday",
      "event": "Labor Day"
   },


….]
```

2. Housing

```
[
   {
      "address": "......",
      "name": "Name of housing",
      "occupancy": [
         "Four-person",
         "Eight-person",
         ………..
      ],
      "overview": "description",
      "room_type": [
         {
            "price": "$$$",
            "type": "type1"
         },
         {
            "price": "$$$",
            "type": "type2"
         },
      …...
```

```
    ],
    "sp_notes": [
        "Air Conditioning",
        "Customer Service Center",
        ……..
    ],
    "type": "Undergraduate"
}
```

3. Dinning

For dinning following attributes would to present if all of them are available.

```
{
Name: " " ,
Logo: " " ,
Address: " ",
Phone: " ",
Inf: " ",
Open hours: " ",
Type: " ",
Keywords: " "
}
```

JSON-LD

The data from the json format is converted to a uniform json ld across all universities, also giving an ease to save them. Also we have converted to a form that is easy to upload to elastic search.

1. Calendar

{"@context":

{"@vocab": "http://schema.org/",

"tags": "http://example.org/tags"},

 "@type": "Event",

"name": "Axess opens for course enrollment.",

 "DayOfWeek": "Wed",

 "tags": "stanford cal",

 "startDate": "August 1"}

2. Housing

{ "@context":
{"@vocab": "http://schema.org/",
"ApartmentType": "http://example.org/ApartmentType",
"ApartmentCategory": "http://example.org/ApartmentCategory",
"sp_notes": "http://example.org/sp_notes",
"tags": "http://example.org/tags"},
"@type": "Apartment",
"name": "Bel-Air ",
"address": "1124 W. 29TH ST.",
"occupancy": ["One family"],
"description": "Bel-Air is a\u00a0small apartment building with one- and

4

two-bedroom\u00a0for married students and families. The one-bedroom units
are a great option for married couples with one child. The two-bedroom
units are reserved for families with two or more children and are in high
demand. Both room types feature recently remodeled kitchens and bathrooms
and living room air-conditioners.\r\n\r\nAll apartments have full kitchens
and bathrooms and an air-conditioner in the living room.",
"sp_notes": ["12-Month Contract", "Air Conditioning", "Parking", "USC Bus
Stop"],
"tags": "usc housing Family",
"ApartmentCategory": "Family Housing",
"ApartmentType": [{"price": "$1,325/Month", "type": "One-bedroom"},
{"price": "$1,430/Month", "type": "Large one-bedroom"},
{"price": "$1,655/Month", "type": "Two-bedroom"}]}

3. Dinning

{
"@context": {"@vocab": "http://schema.org", "tags":
"http://schema.org/tags"},
"name": "The Cafe at the\u00a0Arrillaga\u00a0Alumni Center",
"Logo":
"https://rde-stanford-edu.s3.amazonaws.com/Hospitality/Images/alumniCafe-h
eader2017-110717_0_0.jpg",
"phone": " General Manager(650) 725-9512 | akish@stanford.edu",
"tag": "Stanford dining",
"@type": "Restaurant",
"description": "If you havent tried A\u00e7a\u00ed. the superfoodberry
from the Amazo. nows the time.In our new A\u00e7a\u00edbowlsthe organic
berry is servedblended. then sprinkledwith fruit. granola and your
favorite toppings. Delicious andloaded with antioxidants. its a great way
to start your day. Other savory breakfast items include Smashed Avo on
Toast. What better way to warm up during the lunch hour than with our
Vietnamese pho. rice noodle soup with sliced beef. Our popular Grain

Bowls are a one-dish lunchtime adventure layered with different flavors. We also offer lunch and dinner paninis. signature sandwiches. and grab & go selections.. Join us for Happy Hour 5-7 pm Monday through Friday for sangria. premier wines anddraft beers.. Don\\t forget to join our Coffee Club\u2014buy 10 specialty coffees. and your 11th is on the house!. The Alumni Caf\u00e9 is an ideal setting for meetings. parties and events any time of year. To learn more about our off-site catering services. please see our catering menu. Click here to access our online catering request form.. Ani Kish",
"openingHours": " Monday - Friday: 7:30 a.m. \u2013 7:00 p.m.Breakfast: 7:30 a.m. \u2013 10:00 a.m.Bar service: 3:00 p.m. \u2013 7:00 p.m.Arrillaga Alumni Center \u2013 326 Galvez St.",
"hasMenu": "https://rde-stanford-edu.s3.amazonaws.com/Hospitality/Catering/The%20Cafe%20catering%20menu%20072318_KR.pdf", "kind": "Retail_Caf\u00e9s",
"keywords": ["Assorted", "Service", "Salad", "Water", "Coffee"]}

Ontology with Protege


```
▼─ ● owl:Thing
   ▼─ ● Domain_enity
      ▼─ ● Independent_entity
         ▼─ ● University
               ● MIT
               ● Stanford
               ● USC
         ▼─ ● University_Calander
               ● calander_date
               ● calander_day
               ● calander_event
         ▼─ ● University_Dinning
               ● dinning_address
               ● dinning_hours
               ● dinning_information
               ● dinning_logo
               ● dinning_menu
               ● dinning_name
               ● dinning_phone
               ● dinning_type
         ▼─ ● University_Housing
               ● housing_address
               ● housing_degree
               ● housing_name
            ▼─ ● housing_occupancy
                  ● housing_double
                  ● housing_quad
                  ● housing_single
                  ● housing_triple
               ● housing_overview
            ▼─ ● housing_roomType
               ▼─ ● housing_room
                     ● housing_price
                     ● housing_type
               ● housing_specialNotes
      ▼─ ● Value
         ▼─ ● Expense_level
```

Uploading data to Elasticsearch

Than we uploaded the data to elasticsearch using following curl command.

curl -H 'Content-Type: application/x-ndjson' -XPOST 'localhost:9200/project3/all/_bulk?pretty'
--data-binary '@housingFinal.json'

Algorithm about string match and similarity sorting

1. String match:

```javascript
app.get('/search', function (req, res){
  var body = {
    size: 200,
    from: 0,
    query: {
      bool: {
        should:
        [
          {match: {tags: req.query['q']}},
          {match: {startDate: req.query['q']}},
          {match: {endDate: req.query['q']}},
          {match: {DayOfWeek: req.query['q']}},
          {match: {name: req.query['q']}},
          {match: {address: req.query['q']}},
          {match: {occupancy: req.query['q']}},
          {match: {openingHours: req.query['q']}},
          {match: {description: req.query['q']}},
          {match: {keywords: req.query['q']}},
          {match: {kind: req.query['q']}},
          {match: {sp_notes: req.query['q']}},
          {match: {ApartmentType: req.query['q']}},
          {match: {ApartmentCategory: req.query['q']}}
        ]
      }
    }
  }
```

We use the Should-Match to get the data matching the query which the query is part of one of our terms.

2. Similarity sorting

```
var list = []
for(var i in data){
  var values = $.map(data[i]["_source"],function(value,key){return value});
  var simi = []
  for (var j in values){
    cur_sim = Math.round(similarity(query,values[j].toString())*10000)/100;
    simi.push(cur_sim)
    }
  var s_value = simi.sort((a,b) => a>b).slice(0,5);
  var s_sum = s_value.reduce(function(a,b){return a+b;},0);
  list.push({key:data[i],value:s_sum});
  }

list.sort((a, b) => (a.value < b.value) ? 1 : -1)
// list_sort = Object.keys(list).sort(function(a,b){return list[b]-list[a]});
console.log(this.results.length);
len = list.length
```

```
function similarity(s1, s2) {
    var longer = s1;
    var shorter = s2;
    if (s1.length < s2.length) {
      longer = s2;
      shorter = s1;
    }
    var longerLength = longer.length;
    if (longerLength === 0) {
      return 1.0;
    }
    return (longerLength - editDistance(longer, shorter)) / parseFloat(longerLength);
  }

function editDistance(s1, s2) {
  s1 = s1.toLowerCase();
  s2 = s2.toLowerCase();

  var costs = new Array();
  for (var i = 0; i <= s1.length; i++) {
    var lastValue = i;
    for (var j = 0; j <= s2.length; j++) {
      if (i == 0)
        costs[j] = j;
      else {
        if (j > 0) {
          var newValue = costs[j - 1];
          if (s1.charAt(i - 1) != s2.charAt(j - 1))
            newValue = Math.min(Math.min(newValue, lastValue),
              costs[j]) + 1;
          costs[j - 1] = lastValue;
          lastValue = newValue;
        }
      }
    }
```
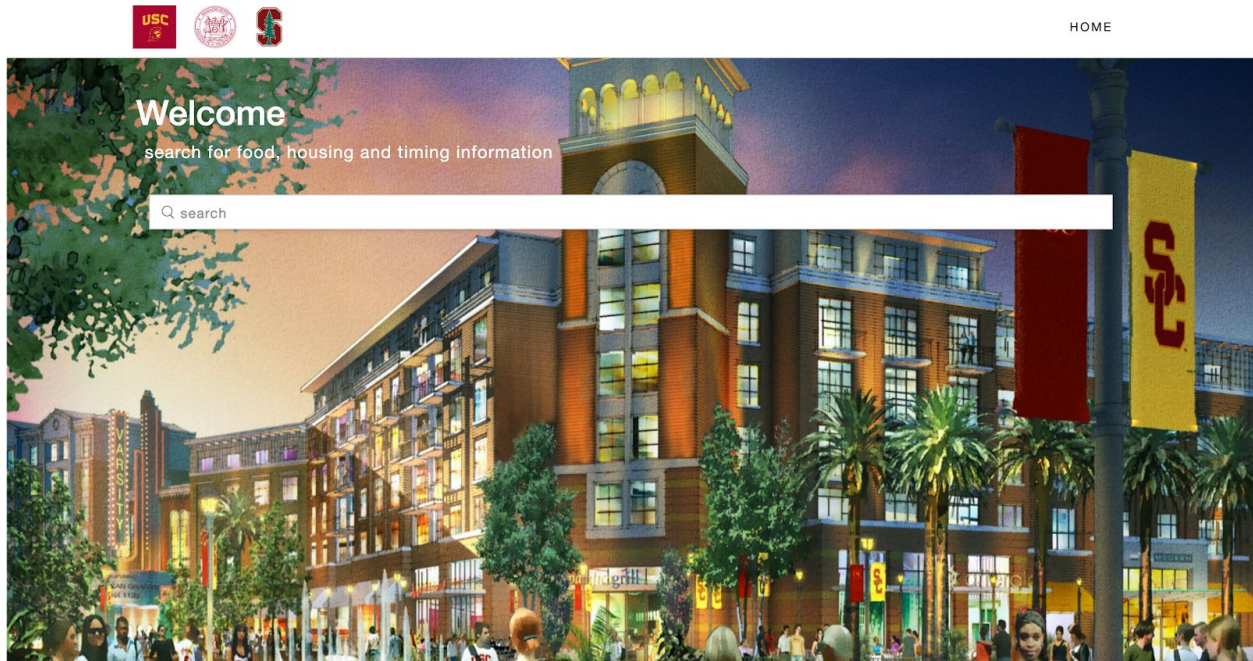
We used the edit distance as the similarity score for each pair between the query and the terms of each row of data, and we get the 5 (it can be changed if needed)

highest values for each row of data and sum them up as the similarity between the query and this row. Then we sort them and show them out in the research result.

User Interface and Searching result

Note: Calendar (baby Green), Housing (baby red), Dining (baby blue)

Q usc

March 10, March 17, Sun-Sun	December 13 – January 6, undefined, Thu-Sun	July 4, July 5, Thu-Fri	January 3, January 4, Thu-Fri
Event: Spring Recess, **tags:** usc cal.	**Event:** Winter Recess, **tags:** usc cal.	**Event:** Independence Day, **tags:** usc cal.	**Event:** Open Registration, **tags:** usc cal.

November 21, November 25, Wed-Sun

Event: Thanksgiving Holiday, **tags:** usc cal.

University Club

705 W. 34th St. Los Angeles, CA 90089
213.740.2030 universityclub.usc.edu

keywords: "Not Found",
Description: The University

🔍 mit

June 7, undefined, Friday

Event: Commencement,
tags: mit cal.

October 8, 9, undefined, Mon, Tues

Event: Columbus Day — vacation,
tags: mit cal.

April 15, 16, undefined, Mon, Tues

Event: Patriots' Day — vacation,
tags: mit cal.

March 25, March 29, Mon-Fri

Event: Spring vacation,
tags: mit cal.

January 7, undefined, Monday

Event: IAP begins,
tags: mit cal.

🔍 stanford

June 15, undefined, Sat

Event: Senior Class Day.,
tags: stanford cal.

Branner Dining

undefined
Not Found

keywords: "Not Found",
Description: Brannersintimate dining room and beautiful estate-style kitchen andserverygreet students five days a week. Seven different food stations offer a variety of cuisine options. and special consideration is

December 15, undefined, Sat, noon

Event: Undergraduate housing closes for Winter Break; see,
tags: stanford cal.

March 18, March 22, Mon-Fri

Event: End-Quarter examinations.,
tags: stanford cal.

December 10, December 14, Mon-Fri

Event: End-Quarter examinations.,
tags: stanford cal.

🔍 cal

February 1, undefined, Friday

Event: IAP ends,
tags: mit cal.

June 15, undefined, Sat

Event: Baccalaureate Saturday.,
tags: stanford cal.

September 3, undefined, Monday

Event: Labor Day,
tags: mit cal.

June 15, undefined, Sat

Event: Senior Class Day.,
tags: stanford cal.

January 7, undefined, Monday

Event: IAP begins,
tags: mit cal.

November 12, undefined, Monday

Event: Veterans Day — holiday,
tags: mit cal.

Traditions

3551 Trousdale Pkwy Los Angeles, CA 90089
213.821.3445
www.usctraditions.com

keywords: "Not Found",
Description: Located within the impressive Ronald Tutor Campus Center, Traditions Bar & Grill adds a sophisticated twist to the pub feel and redefines what on-campus fun is all about. Watch a big game, share a

MIT Faculty Dining

undefined
undefined

keywords: "Not Found",
Description: For reservations, please contact Sally Hennemuth at ,
tags: mit dining.

Dunkin' Donuts

undefined
undefined

keywords: "Not Found",
Description: ,
tags: mit dining.

Shinkansen

undefined
undefined

keywords: "Not Found",
Description: Shinkansen offers Japanese cusine in the Lobdell Food Court.,
tags: mit dining.

Cafe Spice

undefined
undefined

Zigo

undefined
undefined

Cambridge Grill

undefined
undefined

New House

471- 476 Memorial Drive Building W70 Cambridge, MA 02139 (617) 253-6561

Wehave: [{ "price": "$5,895/semester", "type": "single" }, { "price": "$5,215/semester", "type": "double" }],
Description: furnished coed residence for undergraduate students opened in 1975,
tags: MIT housing Undergraduate.

Next House

500 Memorial Drive Building W71 Cambridge, MA 02139 (617) 253-8761

Wehave: [{ "price": "$5,895/semester", "type": "single" }, { "price": "$4,865/semester", "type": "double" }, { "price": "$4,355/semester", "type": "triple" }],
Description: furnished coed residence for undergraduate students opened in 1982,
tags: MIT housing

Eastgate

Building E55 60 Wadsworth Street Cambridge, MA 02142 (617) 452-7463

Wehave: [{ "price": "$1,801/month", "type": "1-bedroom corner apt. " }, { "price": "$1,672/month", "type": "1-bedroom center apt." }, { "price": "$2,148/month", "type": "2-bedroom apt." }],
Description: One of two family residences at MIT, Eastgate Apartments (E55) is

Westgate

Building W85 540 Memorial Drive Cambridge, MA 02139 (617) 253-5146

Wehave: [{ "price": "$2,017/month", "type": "2-bedroom apt. " }, { "price": "$1,787/month", "type": "1-bedroom apt." }, { "price": "$1,517/month", "type": "efficiency apt." }],
Description: One of two family residences at MIT, Westgate Apartments (W85) is located on the west end of

air condition house

Next House

500 Memorial Drive Building W71 Cambridge, MA 02139 (617) 253-8761

Wehave: [{ "price": "$5,895/semester", "type": "single" }, { "price": "$4,865/semester", "type": "double" }, { "price": "$4,355/semester", "type": "triple" }],

Description: furnished coed residence for undergraduate students opened in 1982,

tags: MIT housing

MacGregor House

Building W61 450 Memorial Drive Cambridge, MA 02139 (617) 253-1461

Wehave: [{ "price": "$5,505/semester", "type": "single" }, { "price": "$4,865/semester", "type": "double" }],

Description: furnished coed residence for undergraduate students opened in 1970,

tags: MIT housing Undergraduate.

Annenberg House

711 W. 27TH ST.

Wehave: [{ "price": "$5,640/Semester", "type": "Two-bedroom" }, { "price": "$4,460/Semester", "type": "Two-bedroom" }],

Description: This three-story apartment building is the place for students looking for an active and vibrant community on the leading edge of communications and

Ashdown House

Building NW35 235 Albany Street Cambridge, MA 02139 (617) 253-2961

Wehave: [{ "price": "$1,022/person", "type": "3-bedroom suite" }, { "price": "$1,165/month", "type": "3-bedroom apt. " }, { "price": "$1,401/month", "type": "2-bedroom apt." }, { "price": "$1,597/month", "type": "efficiency apt." }],

Description: Located in the northwest corridor, Ashdown

mit dining

MIT Faculty Dining

undefined
undefined

keywords: "Not Found",
Description: For reservations, please contact Sally Hennemuth at ,
tags: mit dining.

Dunkin' Donuts

undefined
undefined

keywords: "Not Found",
Description: ,
tags: mit dining.

Cambridge Grill

undefined
undefined

keywords: "Not Found",
Description: Enjoy hot-off-the-grill specialties, stone-oven pizza, salads, and more!,
tags: mit dining.

Cafe Spice

undefined
undefined

keywords: "Not Found",
Description: Cafe Spice serves Indian cuisine favorites including appetizers and kathi rolls, as well as chicken tikka masala and chicken vindaloo.,
tags: mit dining.

Q stanford cal

June 15, undefined,
Sat

Event: Senior Class Day.,
tags: stanford cal.

November 30, undefined,
Fri

Event: Last day of Law classes. See the,
tags: stanford cal.

August 15, undefined,
Thu

Event: Last day of classes.,
tags: stanford cal.

June 5, undefined,
Wed

Event: Last day of classes.,
tags: stanford cal.

February 10, undefined,
Sun

Event: Axess opens for course enrollment.,
tags: stanford cal.

April 14, undefined,
Sun

Event: Axess opens for course enrollment.,
tags: stanford cal.

Q last day of classes

May 16, undefined,
Thursday

Event: Last day of classes,
tags: mit cal.

August 15, undefined,
Thu

Event: Last day of classes.,
tags: stanford cal.

June 5, undefined,
Wed

Event: Last day of classes.,
tags: stanford cal.

December 12, undefined,
Wednesday

Event: Last day of classes,
tags: mit cal.

February 5, undefined,
Tuesday

Event: First day of classes,
tags: mit cal.

September 5, undefined,
Wednesday

Event: First day of classes,
tags: mit cal.

May 30, undefined,
Thu

Event: Last day of Law classes. See the,
tags: stanford cal.

March 11, undefined,
Mon

Event: Last day of Law classes. See the,
tags: stanford cal.

Q Figueroa dining

MIT Faculty Dining

undefined
undefined

keywords: "Not Found",
Description: For reservations, please contact Sally Hennemuth at ,
tags: mit dining.

Gerhard Dining

undefined
Not Found

keywords: "Not Found",
Description: With its renovation in 2015. Gerhard Casper Dining offers fantastic grilled items. frozen yogurt. signature salads. and daily pizza offerings. Gerhard Casper Dining

---

Wilbur Dining

undefined
Not Found

keywords: "Not Found",
Description: Stanfords largest residential dining complex. Wilbur Dining serves residents of Arroyo. Cedro. Junipero. Okada. Otero. Rinconada. Soto. and Trancos. Its spacious servery

---

Ricker Dining

undefined
Not Found

keywords: "Not Found",
Description: The first university facility of its kind in the natio. Ricker Dining is Stanfords highly lauded peanut and tree nut-sensitive dining environment. In recognition

---

🔍 water tea |

Literatea

Not Found
213.821.4261

keywords: "Not Found",
Description: The perfect compliment to a quiet afternoon with books is a steady supply of good tea or coffee. Literatea has plenty of both, along with a compliment of healthy grab & go items to make your study hours more pleasant

The Axe & Palm Cafe at Old Union

undefined
General Manager(650) 721-6982 | mjordan1@stanford.edu

keywords: ["Stanford", "R&DE", "Café", "Palm", "Catering"],
Description: Old Union – 520 Lasuen Mall. From the Big Game to late-night dining. the Axe & Palm tradition lives on! Enjoy with

Forbes Cafe

undefined
General Manager(650) 724-7969jflores2@stanford.edu

keywords: ["Assorted", "Water", "Tea", "Regular", "Salad"],
Description: Forbes Family Café in now open for dinner Monday – Thursday 5:00 p.m. – 7:30 p.m.. Grad students can use Graduate Meal Plan swipes for

The Cafe at the Arrillaga Alumni Center

undefined
General Manager(650) 725-9512 | akish@stanford.edu

keywords: ["Assorted", "Service", "Salad", "Water", "Coffee"],
Description: If you havent tried Açaí. the superfoodberry from the Amazo. nows the time.In our

🔍 salad

SLACafe

undefined
Service Relations Manager(650) 353-1610 | botellos@stanford.edu

keywords: ["Stanford", "Chicken", "R&DE", "Café", "Cheese"],
Description: Residential & Dining Enterprises' Stanford Hospitality & Auxiliaries welcomes you totheSLAC community's newestrestaurant. Visit the

Forbes Family Cafe

undefined
undefined

keywords: ["Gluten-Containing", "York", "Style", "Fat,value0,unitg,saturatedFatConten "Ingredients,ordered_cor_icon0002-0004id4,labelVegan,0003-0009id9,labelMade"],
Description: Visit the Forbes Family Cafe in MITs Ray and Maria Stata Center for Computer Information

100 Main Marketplace

undefined
undefined

keywords: "Not Found",
Description: Located in the Sloan School of Management, 100 Main Marketplace serves breakfast and lunch. The menu includes made-to-order hot breakfast, panini sandwiches, sushi, soup, and daily market features

Pacific Street Cafe

undefined
undefined

keywords: ["Home", "Hot", "Us", "Contact", "Menu"],
Description: In the morning, hot pastries and freshly-brewed coffee are ready when the cafe opens, or stop by at lunch for a made-to-order sandwich, one of our daily specials, or a cup of soup and salad. Pacific

🔍 salad in usc|

100 Main Marketplace

undefined
undefined

keywords: "Not Found",
Description: Located in the Sloan School of Management, 100 Main Marketplace serves breakfast and lunch. The menu includes made-to-order hot breakfast, panini sandwiches, sushi, soup,

Forbes Family Cafe

undefined
undefined

keywords: ["Gluten-Containing", "York", "Style", "Fat,value0,unitg,saturatedFatConten "Ingredients,ordered_cor_icon0002-0004id4,labelVegan,0003-0009id9,labelMade"],
Description: Visit the Forbes Family Cafe in MITs Ray and Maria Stata Center

Pacific Street Cafe

undefined
undefined

keywords: ["Home", "Hot", "Us", "Contact", "Menu"],
Description: In the morning, hot pastries and freshly-brewed coffee are ready when the cafe opens, or stop by at lunch for a made-to-order sandwich, one of our daily specials, or a cup

Rebecca's Cafe at Walker Memorial

undefined
undefined

keywords: "Not Found",
Description: Choose your own hearty bowl or wrap or choose from signature selections of grab & go salads, sandwiches, snacks, and dessert made fresh in our local bakery!Also

Fairmont

2629 PORTLAND ST.

Wehave: [{ "price": "$1,050/Month", "type": "Bachelor" }, { "price": "$1,215/Month", "type": "Studio" }, { "price": "$840/Month", "type": "One-bedroom" }, { "price": "$635/Month", "type": "Two-bedroom" }],

Description: Fairmont is a favorite choice for graduate

Centennial

2390 PORTLAND ST.

Wehave: [{ "price": "$1,265/Month", "type": "Two-bedroom" }],

Description: Centennial is an off-campus apartment building for graduate students, home to the special interest community for occupational therapy students. Two floors are reserved for residents in the

Twin Palms

2635 PORTLAND ST.

Wehave: [{ "price": "$1,050/Month", "type": "Bachelor" }, { "price": "$1,215/Month", "type": "Studio" }, { "price": "$1,320/Month", "type": "One-bedroom" }, { "price": "$1,110/Month", "type": "Two-bedroom with double bath" }],

Description: Always one of

Parkside Apartments

3730 S. MCCLINTOCK AVE.

Wehave: [{ "price": "$6,650/Semester", "type": "Studio" }, { "price": "$4,610/Semester", "type": "Studio" }, { "price": "$4,900/Semester", "type": "Large studio" }, { "price": "$5,685/Semester", "type": "One-bedroom" }, { "price": "$5,900/Semester", "type": "Two-bedroom single

17

Extra: PDF data extracting

Example PDF:

https://rde-stanford-edu.s3.amazonaws.com/Hospitality/Catering/The%20Cafe%20catering%20menu%20072318_KR.pdf

Alumni Café Catering Menu 2018-19

Breakfast Packages (10 people Min)
All Breakfast Buffet Packages include freshly brewed Starbucks Coffee and Tea Services

Faculty Breakfast $7.50/person
- *Assorted Freshly Baked Bagels (Butter, Jellies, Plain and light Cream Cheese)*
- *Seasonal Whole Fruit*
- *Complimentary Water Service*
- *Fresh Brewed Starbucks Coffee and Tea Service (Sugar, Sweetener and Half & Half) (Decaf upon request)*

Healthy Start Breakfast $8.00/person
- *Organic Granola with yogurt, Served with 2% milk*
- *Hard Boiled Eggs*
- *Seasonal Whole Fruit*
- *Fresh Squeezed Orange Juice*
- *Complimentary Water Service*
- *Fresh Brewed Starbucks Coffee and Tea Service (Sugar, Sweetener and Half & Half) (Decaf upon request)*

Alumni Continental Breakfast $9.50/person
- *Assorted Fresh Baked Breakfast Pastries*
- *Fresh Sliced Fruit Platter*
- *Hard Boiled Eggs*
- *Fresh Squeezed Orange Juice*
- *Complimentary Water Service*
- *Fresh Brewed Starbucks Coffee and Tea Service (Sugar, Sweetener and Half & Half) (Decaf upon request)*

Hot Breakfast $18.50/person
- *Scrambled Eggs,*
- *Sausage or Bacon*
- *Cottage Fries Potatoes,*
- *Fruit Bowl*
- *Assorted Freshly Baked Scones*
- *Fresh Brewed Starbucks Coffee and Tea Service (Sugar, Sweetener and Half & Half) (Decaf*

We used the python library -- PyPDF2 to transform it to text and usc nltk to tag every word, so we can get the hot words from this restaurant like salad, chicken, etc. We finished this procession for all restaurants which has PDF format menu.

```python
with io.BytesIO(response.content) as open_pdf_file:
    read_pdf = PyPDF2.PdfFileReader(open_pdf_file)
    for i in range(read_pdf.numPages):
        page = read_pdf.getPage(i)
        text = page.extractText()
        words += [i.strip(" ").replace("\n","") for i in text.split(" ") if i.strip(" ").replace("\n","") != '']
# tagged = nltk.pos_tag(words)
counter = Counter(words)
keywords = [i[0] for i in counter.most_common(30)]
keywords = [i[0] for i in nltk.pos_tag(keywords) if i[1] == 'NNP' and sum([a not in '%^&*$#@!\'\";.,=+-_()' for a in i[0]]) >= 2]
if len(keywords) >= 5:
    d['keywords'] = keywords[:5]
else:
    d['keywords'] = 'Not Found'
```