

```
In [1]: # pwd
Out[1]: '/Users/sathishrajendiran/ist664-nlp'
```

NLP Homework 2:

Contact Finder: Using Regular Expressions to find email addresses and phone numbers

- Name: Sathish Kumar Rajendiran
- Homework: 2| Contact Finder
- Date: 11/08/2020
- Class: 2020-0930 IST 664- Natural Language Processing

Contents Explored

1) Content Finder Regular expression on .edu email addresses

2) Content Finder Regular expression on .com email addresses

3) Content Finder Regular expression on phone numbers

Objective

Compare the result with Gold Data set

1) Reduce the number of False Positive

2) Reduce the number of False Negative

3) Increase the number of True Positive

4) Challenges

5) Next steps

Sources

1. Data files

| Name | Date Modified | Size | Kind |
|-----------------------|----------------------|-----------|---------------|
| data | Today at 1:34 PM | -- | Folder |
| dev | Yesterday at 2:50 PM | -- | Folder |
| aliken | Yesterday at 2:49 PM | 9 KB | TextEdit |
| ashishg | Yesterday at 2:49 PM | 6 KB | TextEdit |
| balaji | Yesterday at 2:49 PM | 1 KB | TextEdit |
| bgired | Yesterday at 2:49 PM | 14 KB | TextEdit |
| cheriton | Yesterday at 2:49 PM | 1 KB | TextEdit |
| christos | Yesterday at 2:49 PM | 5 KB | TextEdit |
| dabo | Yesterday at 2:49 PM | 7 KB | TextEdit |
| dwh | Yesterday at 2:49 PM | 34 bytes | TextEdit |
| dm | Yesterday at 2:49 PM | 2 KB | TextEdit |
| engler | Yesterday at 2:49 PM | 25 KB | TextEdit |
| eroberts | Yesterday at 2:49 PM | 1 KB | TextEdit |
| fedkiw | Yesterday at 2:49 PM | 60 KB | TextEdit |
| hager | Yesterday at 2:49 PM | 10 KB | TextEdit |
| hanrahan | Yesterday at 2:49 PM | 8 KB | TextEdit |
| horowitz | Yesterday at 2:49 PM | 4 KB | TextEdit |
| jks | Yesterday at 2:49 PM | 7 KB | TextEdit |
| jurafsky | Yesterday at 2:49 PM | 9 KB | TextEdit |
| jure | Yesterday at 2:49 PM | 295 bytes | TextEdit |
| knuth | Yesterday at 2:49 PM | 2 KB | TextEdit |
| koller | Yesterday at 2:50 PM | 7 KB | TextEdit |
| kosecka | Yesterday at 2:49 PM | 7 KB | TextEdit |
| kunle | Yesterday at 2:49 PM | 14 KB | TextEdit |
| lam | Yesterday at 2:49 PM | 18 KB | TextEdit |
| latombe | Yesterday at 2:49 PM | 35 KB | TextEdit |
| levoy | Yesterday at 2:49 PM | 9 KB | TextEdit |
| manning | Yesterday at 2:49 PM | 12 KB | TextEdit |
| nass | Yesterday at 2:49 PM | 7 KB | TextEdit |
| nick | Yesterday at 2:49 PM | 6 KB | TextEdit |
| nickm | Yesterday at 2:49 PM | 100 bytes | TextEdit |
| ok | Yesterday at 2:49 PM | 5 KB | TextEdit |
| oustter | Yesterday at 2:49 PM | 6 KB | TextEdit |
| pal | Yesterday at 2:49 PM | 12 KB | TextEdit |
| plotkin | Yesterday at 2:49 PM | 293 bytes | TextEdit |
| psyoung | Yesterday at 2:49 PM | 2 KB | TextEdit |
| rajeve | Yesterday at 2:49 PM | 16 KB | TextEdit |
| rinard | Yesterday at 2:49 PM | 7 KB | TextEdit |
| serafim | Yesterday at 2:49 PM | 7 KB | TextEdit |
| shoham | Yesterday at 2:50 PM | 8 KB | TextEdit |
| subh | Yesterday at 2:49 PM | 12 KB | TextEdit |
| thm | Yesterday at 2:49 PM | 1 KB | TextEdit |
| tim | Yesterday at 2:49 PM | 5 KB | TextEdit |
| ullman | Yesterday at 2:49 PM | 8 KB | TextEdit |
| vladien | Yesterday at 2:49 PM | 7 KB | TextEdit |
| widom | Yesterday at 2:49 PM | 16 KB | TextEdit |
| zelenski | Yesterday at 2:49 PM | 4 KB | TextEdit |
| zlm | Yesterday at 2:49 PM | 10 KB | TextEdit |
| devGOLD | Yesterday at 2:49 PM | 3 KB | TextEdit |
| ContactFinder.base.py | Yesterday at 2:46 PM | 7 KB | Python Script |
| ContactFinder.py | Today at 1:27 PM | 11 KB | Python Script |

Loading and Cleaning the Data

- Download the ContentFinder Raw files and python scripts.
- Console setup

```
(base) Sathishs-MacBook-Pro:ContactFinder sathishrajendiran$ pwd
/Users/sathishrajendiran/ist664-nlp/HW/ContactFinder
(base) Sathishs-MacBook-Pro:ContactFinder sathishrajendiran$ ls -ltr
total 96
-rw-r--r--  1 sathishrajendiran  staff   6748 Nov  8 14:46 ContactFinder.base.py
-rw-r--r--  1 sathishrajendiran  staff  11861 Nov  9 13:27 ContactFinder.py
-rw-r--r--  1 sathishrajendiran  staff   2596 Nov  9 13:28 untitled.ipynb
drwxr-xr-x  2 sathishrajendiran  staff    192 Nov  9 13:34 data
(base) Sathishs-MacBook-Pro:ContactFinder sathishrajendiran$

# email .edu patterns

# each regular expression pattern should have exactly two sets of parentheses.
# the first parenthesis should be around the someone part
# the second parenthesis should be around the somewhere part
# in an email address whose standard form is someone@somewhere.edu
epatterns = []
epatterns.append('([A-Za-z.]+)@([A-Za-z.]+\.)\.edu')
epatterns.append('([A-Za-z.]+\.)@([A-Za-z.]+\.)\.edu')

# email .com patterns

# each regular expression pattern should have exactly two sets of parentheses.
# the first parenthesis should be around the someone part
# the second parenthesis should be around the somewhere part
# in an email address whose standard form is someone@somewhere.com
ecompatterns = []
ecompatterns.append('([A-Za-z.]+)@([A-Za-z.]+\.)\.com')

# Regular expression replacement on for overall

line = re.sub("<at symbol> ", "", line) #replace <at symbol> with 
line = re.sub("<lt> ", "", line) #replace <lt> with "
line = re.sub("<gt> ", "", line) #replace <gt> with "
line = re.sub(" at <!-- die!!--> ", "", line) #replace <!-- die!!--> with "
line = re.sub(" at <!-- spam pigs!!--> dot <!-- die!!--> ", "", line) #replace <!-- spam pigs!!--> with ""
line = re.sub("#<#40> ", "", line) #replace #<#40> with @
line = re.sub("\<followed by &ldquo>;", "", line) #replace (followed by &ldquo; with ""
line = re.sub("<rdquo>)", "", line) #replace (followed by &ldquo; with ""
line = re.sub("\<followed by \",", "", line) #replace (followed by " with ""

line = re.sub("\\""", "", line) #replace " with ""
line = re.sub(" AT ", "", line) #replace 'at with @
line = re.sub(" DOT ", "", line) #replace 'dot with .
line = re.sub("<del> ", "", line) #replace <del> with "
line = re.sub(" @ ", "", line) #replace @ with @
line = re.sub(" WHERE ", "", line) #replace 'WHERE with @
line = re.sub(" DOM ", "", line) #replace 'DOM with @
line = re.sub(" dt ", "", line) #replace 'dt with @

# Regular expression replacement on for Phone numbers

line = re.sub("\[650\] ", "650-", line) #replace [650] with 650-

pline = re.sub(" ", "-", line) #replace  with -
pline = re.sub("(", "", line) #replace ( with blank
pline = re.sub(")", "", line) #replace ) with blank
pline = re.sub("\\""", "", line) #replace " with "
pline = re.sub("<BR>", "", line) #replace <BR> with space
pline = re.sub("<br />", "", line) #replace <br /> with blank

pline2 = re.sub("\+1 650 ", "650-", pline) #replace +1 650 with 650-
pline2 = re.sub("": "\(", "", pline2) #replace : \ ( with blank
pline2 = re.sub(" ", "-", pline2) # replace # ) with -

# Regular expression replacement on for .com email

ecomline = re.sub(" at ", "", ecomline) #replace at with @
ecomline = re.sub(" dt ", "", ecomline) #replace dt with .

# Regular expression replacement on for Phone numbers

line = re.sub("\[650\] ", "650-", line) #replace [650] with 650-

pline = re.sub(" ", "-", line) #replace  with -
pline = re.sub("(", "", line) #replace ( with blank
pline = re.sub(")", "", line) #replace ) with blank
pline = re.sub("\\""", "", line) #replace " with "
pline = re.sub("<BR>", "", line) #replace <BR> with space
pline = re.sub("<br />", "", line) #replace <br /> with blank

pline2 = re.sub("\+1 650 ", "650-", pline) #replace +1 650 with 650-
pline2 = re.sub("": "\(", "", pline2) #replace : \ ( with blank
pline2 = re.sub(" ", "-", pline2) # replace # ) with -

# email pattern with .edu list
for epat in epatterns:
    # each epat has 2 sets of parentheses so each match will have 2 items in a list
    matches = re.findall(epat, eline, re.IGNORECASE) #ignore case
    for m in matches:
        # string formatting operator % takes elements of list m
        # and inserts them in place of each %s in the result string
        # email has form someone@somewhere.edu
        # email = '%s@%s.edu' % m
        email = '{}@{}.edu'.format(m[0], m[1])
        res.append((name, 'e', email))

# email pattern with .com list
for ecompat in ecompatterns:
    # each ecompat has 2 sets of parentheses so each match will have 2 items in a list
    matches = re.findall(ecompat, ecomline, re.IGNORECASE) #ignore case
    for m in matches:
        # string formatting operator % takes elements of list m
        # and inserts them in place of each %s in the result string
        # email has form someone@somewhere.edu
        # email = '%s@%s.edu' % m
        email = '{}@{}.com'.format(m[0], m[1])
        res.append((name, 'e', email))

# phone pattern list
for ppat in ppatterns:
    # each ppat has 3 sets of parentheses so each match will have 3 items in a list
    matches = re.findall(ppat, pline3)
    for m in matches:
        # phone number has form areacode-exchange-number
        #phone = '%s-%s-%s' % m
        phone = '{}-{}-{}'.format(m[0], m[1], m[2])
        res.append((name, 'p', phone))
```

Code Execution

- Command line execution of the Python script

```
(base) Sathishs-MacBook-Pro:ContactFinder sathishrajendiran$ python ContactFinder.py
Assuming ContactFinder.py called in directory with data folder
```

True Positives (116):

```
{('ashishg', 'e', 'ashishg@stanford.edu'),
 ('ashishg', 'e', 'rozm@stanford.edu'),
 ('ashishg', 'p', '650-723-1614'),
 ('ashishg', 'p', '650-723-4173'),
 ('ashishg', 'p', '650-814-1478'),
 ('balaji', 'e', 'balaji@stanford.edu'),
 ('bgired', 'e', '650-723-4539'),
 ('bgired', 'p', '650-724-3648'),
 ('bgired', 'p', '650-724-6354'),
 ('cheriton', 'e', 'cheriton@cs.stanford.edu'),
 ('cheriton', 'e', 'uma@cs.stanford.edu'),
 ('cheriton', 'p', '650-723-1131'),
 ('cheriton', 'p', '650-725-3726'),
 ('dabo', 'e', 'dabo@cs.stanford.edu'),
 ('dabo', 'p', '650-725-3897'),
 ('dabo', 'p', '650-725-4671'),
 ('dlwh', 'e', 'dlwh@stanford.edu'),
 ('engler', 'e', 'engler@lcs.mit.edu'),
 ('engler', 'e', 'engler@stanford.edu'),
 ('eroberts', 'e', 'eroberts@cs.stanford.edu'),
 ('eroberts', 'p', '650-723-3642'),
 ('eroberts', 'p', '650-723-6892'),
 ('fedkiw', 'e', 'fedkiw@cs.stanford.edu'),
 ('hager', 'p', '410-516-5521'),
 ('hager', 'p', '410-516-5553'),
 ('hager', 'p', '410-516-8000'),
 ('hanrahan', 'e', 'hanrahan@cs.stanford.edu'),
 ('hanrahan', 'p', '650-723-0033'),
 ('hanrahan', 'p', '650-723-8530'),
 ('horowitz', 'p', '650-725-6707'),
 ('horowitz', 'p', '650-725-3949'),
 ('jks', 'e', 'jks@robotics.stanford.edu'),
 ('kosecka', 'e', 'kosecka@cs.gmu.edu'),
 ('kosecka', 'p', '783-993-1710'),
 ('kosecka', 'p', '783-993-1876'),
 ('kunle', 'e', 'darlene@csl.stanford.edu'),
 ('kunle', 'e', 'kunle@ogun.stanford.edu'),
 ('kunle', 'p', '650-723-1430'),
 ('kunle', 'p', '650-725-3713'),
 ('lam', 'e', 'lam@cs.stanford.edu'),
 ('lam', 'p', '650-725-3714'),
 ('lam', 'p', '650-725-6949'),
 ('latombe', 'e', 'asandra@cs.stanford.edu'),
 ('latombe', 'e', 'latombe@cs.stanford.edu'),
 ('latombe', 'e', 'liliana@cs.stanford.edu'),
 ('latombe', 'p', '650-721-6625'),
 ('latombe', 'p', '650-723-0350'),
 ('latombe', 'p', '650-723-4137'),
 ('latombe', 'p', '650-725-1449'),
 ('levoy', 'e', 'ada@graphics.stanford.edu'),
 ('levoy', 'e', 'melissa@graphics.stanford.edu'),
 ('levoy', 'p', '650-723-0833'),
 ('levoy', 'p', '650-724-6865'),
 ('levoy', 'p', '650-725-3724'),
 ('levoy', 'p', '650-725-4089'),
 ('manning', 'e', 'dbarros@cs.stanford.edu'),
 ('manning', 'e', 'manning@cs.stanford.edu'),
 ('manning', 'p', '650-723-7683'),
 ('manning', 'p', '650-725-1449'),
 ('manning', 'p', '650-725-3358'),
 ('nass', 'e', 'nass@stanford.edu'),
 ('nass', 'p', '650-723-5499'),
 ('nass', 'p', '650-725-2472'),
 ('nick', 'e', 'nick.parlante@cs.stanford.edu'),
 ('nick', 'p', '650-725-4727'),
 ('ouster', 'e', 'ouster@cs.stanford.edu'),
 ('ouster', 'e', 'teresa.lynn@stanford.edu'),
 ('pal', 'e', 'pal@cs.stanford.edu'),
 ('psyoung', 'e', 'patrick.young@stanford.edu'),
 ('rajeve', 'p', '650-723-4671'),
 ('rajeve', 'p', '650-723-4648'),
 ('rajeve', 'p', '650-725-4671'),
 ('rinard', 'e', 'rinard@lcs.mit.edu'),
 ('rinard', 'p', '657-250-1221'),
 ('rinard', 'p', '657-250-6922'),
 ('serafim', 'e', 'serafim@cs.stanford.edu'),
 ('serafim', 'p', '650-723-3334'),
 ('serafim', 'p', '650-725-1449'),
 ('shoham', 'e', 'shoham@stanford.edu'),
 ('shoham', 'p', '650-723-3632'),
 ('shoham', 'p', '650-725-1449'),
 ('subh', 'e', 'subh@stanford.edu'),
 ('subh', 'e', 'uma@cs.stanford.edu'),
 ('subh', 'p', '650-724-1915'),
 ('subh', 'p', '650-725-3724'),
 ('subh', 'p', '650-725-6949'),
 ('thm', 'e', 'pkrokel@stanford.edu'),
 ('thm', 'p', '650-725-3283'),
 ('thm', 'p', '650-725-3636'),
 ('thm', 'p', '650-725-3938'),
 ('tim', 'p', '650-725-0147'),
 ('tim', 'p', '650-725-2348'),
 ('ullman', 'e', 'support@gradience.com'),
 ('ullman', 'e', 'ullman@cs.stanford.edu'),
 ('ullman', 'p', '650-494-8816'),
 ('ullman', 'p', '650-723-2588'),
 ('ullman', 'p', '650-725-6882'),
 ('vladien', 'e', 'vladien@stanford.edu'),
 ('widom', 'e', 'siroker@cs.stanford.edu'),
 ('widom', 'e', 'widom@cs.stanford.edu'),
 ('widom', 'p', '650-723-0872'),
 ('widom', 'p', '650-723-7690'),
 ('zelenski', 'p', '650-725-2588'),
 ('zelenski', 'e', 'zelenski@cs.stanford.edu'),
 ('zelenski', 'p', '650-723-6892'),
 ('zelenski', 'p', '650-725-8556'),
 ('zlm', 'e', 'mann@cs.stanford.edu'),
 ('zlm', 'p', '650-723-4364'),
 ('zlm', 'p', '650-725-4671')}
```

```
(base) Sathishs-MacBook-Pro:ContactFinder sathishrajendiran$
```

```
{('latombe', 'p', '650-721-6625'),
 ('latombe', 'p', '650-723-0350'),
 ('latombe', 'p', '650-723-4137'),
 ('latombe', 'p', '650-725-1449'),
 ('levoy', 'e', 'ada@graphics.stanford.edu'),
 ('levoy', 'e', 'melissa@graphics.stanford.edu'),
 ('levoy', 'p', '650-723-0833'),
 ('levoy', 'p', '650-724-6865'),
 ('levoy', 'p', '650-725-3724'),
 ('levoy', 'p', '650-725-4089'),
 ('manning', 'e', 'dbarros@cs.stanford.edu'),
 ('manning', 'e', 'manning@cs.stanford.edu'),
 ('manning', 'p', '650-723-7683'),
 ('manning', 'p', '650-725-1449'),
 ('manning', 'p', '650-725-3358'),
 ('nass', 'e', 'nass@stanford.edu'),
 ('nass', 'p', '650-723-5499'),
 ('nass', 'p', '650-725-2472'),
 ('nick', 'e', 'nick.parlante@cs.stanford.edu'),
 ('nick', 'p', '650-725-4727'),
 ('ouster', 'e', 'ouster@cs.stanford.edu'),
 ('ouster', 'e', 'teresa.lynn@stanford.edu'),
 ('pal', 'e', 'pal@cs.stanford.edu'),
 ('psyoung', 'e', 'patrick.young@stanford.edu'),
 ('rajeve', 'p', '650-723-4671'),
 ('rajeve', 'p', '650-723-4648'),
 ('rajeve', 'p', '650-725-4671'),
 ('rinard', 'e', 'rinard@lcs.mit.edu'),
 ('rinard', 'p', '657-250-1221'),
 ('rinard', 'p', '657-250-6922'),
 ('serafim', 'e', 'serafim@cs.stanford.edu'),
 ('serafim', 'p', '650-723-3334'),
 ('serafim', 'p', '650-725-1449'),
 ('shoham', 'e', 'shoham@stanford.edu'),
 ('shoham', 'p', '650-723-3632'),
 ('shoham', 'p', '650-725-1449'),
 ('subh', 'e', 'subh@stanford.edu'),
 ('subh', 'e', 'uma@cs.stanford.edu'),
 ('subh', 'p', '650-724-1915'),
 ('subh', 'p', '650-725-3724'),
 ('subh', 'p', '650-725-6949'),
 ('thm', 'e', 'pkrokel@stanford.edu'),
 ('thm', 'p', '650-725-3283'),
 ('thm', 'p', '650-725-3636'),
 ('thm', 'p', '650-725-3938'),
 ('tim', 'p', '650-725-0147'),
 ('tim', 'p', '650-725-2348'),
 ('ullman', 'e', 'support@gradience.com'),
 ('ullman', 'e', 'ullman@cs.stanford.edu'),
 ('ullman', 'p', '650-494-8816'),
 ('ullman', 'p', '650-723-2588'),
 ('ullman', 'p', '650-725-6882'),
 ('vladien', 'e', 'vladien@stanford.edu'),
 ('widom', 'e', 'siroker@cs.stanford.edu'),
 ('widom', 'e', 'widom@cs.stanford.edu'),
 ('widom', 'p', '650-723-0872'),
 ('widom', 'p', '650-723-7690'),
 ('zelenski', 'p', '650-725-2588'),
 ('zelenski', 'e', 'zelenski@cs.stanford.edu'),
 ('zelenski', 'p', '650-723-6892'),
 ('zelenski', 'p', '650-725-8556'),
 ('zlm', 'e', 'mann@cs.stanford.edu'),
 ('zlm', 'p', '650-723-4364'),
 ('zlm', 'p', '650-725-4671')}
```

```
False Positives (1):
{('jurafsky', 'e', 'jurafsky@stanford.edu')}
Summary: tp=116, fp=0, fn=1
```

```
(base) Sathishs-MacBook-Pro:ContactFinder sathishrajendiran$
```

Summary

In this exercise, we have explored 2 major content types

- Email
 - Email category - split into .edu and .com domains.
- Phone numbers
 - Phone numbers - were all assumed to be US number with format as xxx-xxx-xxxx

- Based on the regular expression templates - all 47 files were analyzed for matching email and Phone number patterns. Number of trials depends on the combination of changed needed. Often, I ran into False Positive scenarios, as in image below. However, as we progressed - I split the lines based on email vs phone number to keep the changes not to affect the other expressions.

ex. replacing "-" with "blank space" - did help fix "d-i-w-h-s-t-a-n-f-o-r-d-e-d-u". However, it created false positives dealing with phone numbers. So, I often had to keep the changes to specific patterns unique.

| Trials | True Positive | False Positive | False Negative | comments |
|--------|---------------|----------------|----------------|---|
| 1 | 22 | 0 | 25 | no changes |
| 2 | 41 | 0 | 76 | uncomment telephone patterns |
| 3 | 42 | 0 | 75 | ignore case |
| 4 | 44 | 0 | 73 | replace <at symbol> with @ |
| 5 | 46 | 0 | 71 | replace with @ |
| 6 | 47 | 0 | 70 | (followed by “ |
| 7 | 48 | 0 | 69 | (followed by " |
| 8 | 49 | 0 | 68 | replace <lt> with @ and DOT with . |
| 9 | 53 | 2 | 64 | replace <at> with @ and dot with . |
| 10 | 54 | 2 | 63 | replace " @ " with @ |
| 11 | 55 | 3 | 62 | replace " @ " with @ |
| 12 | 56 | 3 | 61 | Replace WHERE with @ and DOM with . |
| 13 | 59 | 3 | 58 | replace with "" |
| 14 | 60 | 2 | 57 | replace " @ " with @ |
| 15 | 61 | 2 | 56 | replace <!-- die!!--> & <!-- spam pigs!!--> |

| Trials | True Positive | False Positive | False Negative | comments |
|--------|---------------|----------------|----------------|---|
| 1 | 22 | 0 | 95 | no changes |
| 2 | 41 | 0 | 76 | unreplace telephone patterns |
| 3 | 42 | 0 | 75 | ignore case |
| 4 | 44 | 0 | 73 | replace <at symbol> with @ |
| 5 | 46 | 0 | 71 | replace @ with @ |
| 6 | 47 | 0 | 70 | (followed by ` |
| 7 | 48 | 0 | 69 | followed by " |
| 8 | 49 | 0 | 68 | replace AT with @ and DOT with . |
| 9 | 53 | 0 | 64 | replace at with @ and dot with . |
| 10 | 54 | 2 | 63 | replace " with @ |
| 11 | 55 | 3 | 62 | replace with . |
| 12 | 56 | 3 | 61 | Replace WHERE with @ and DOM with . |
| 13 | 59 | 3 | 58 | replace with "" |
| 14 | 60 | 2 | 57 | replace " with @ |
| 15 | 61 | 2 | 56 | replace <!-- del--> & <!-- spam pig!--> |

| Trials | Actual Value | Possible Reason | True Positive | False Positive | % Complete |
|---|--------------------------|---|---------------|----------------|------------|
| {('ashishg', 'p', '650-723-1614'), ('ashishg', 'p', '650-723-4173'), ('ashishg', 'p', '650-814-1478'), ('bgired', 'p', '650-723-3648'), ('bgired', 'p', '650-724-6354'), ('cheriton', 'e', 'uma@cs.stanford.edu'), ('dabo', 'e', 'dabo@cs.stanford.edu'), ('dabo', 'p', '650-725-3897'), ('dabo', 'p', '650-725-4671'), ('dlwh', 'e', 'dlwh@stanford.edu'), ('engler', 'e', 'engler@lcs.mit.edu'), ('engler', 'e', 'engler@stanford.edu'), ('eroberts', 'e', 'eroberts@cs.stanford.edu'), ('eroberts', 'p', '650-723-3642'), ('eroberts', 'p', '650-723-6892'), ('fedkiw', 'e', 'fedkiw@cs.stanford.edu'), ('hager', 'p', '410-516-5521'), ('hager', 'p', '410-516-5553'), ('hager', 'p', '410-516-8000'), ('hanrahan', 'e', 'hanrahan@cs.stanford.edu'), ('hanrahan', 'p', '650-723-0033'), ('hanrahan', 'p', '650-723-8530'), ('horowitz', 'p', '650-725-6707'), ('horowitz', 'p', '650-725-3949'), ('jks', 'e', 'jks@robotics.stanford.edu'), ('kosecka', 'e', 'kosecka@cs.gmu.edu'), ('kosecka', 'p', '783-993-1710'), ('kosecka', 'p', '783-993-1876'), ('kunle', 'e', 'darlene@csl.stanford.edu'), ('kunle', 'e', 'kunle@ogun.stanford.edu'), ('kunle', 'p', '650-723-1430'), ('kunle', 'p', '650-725-3713'), ('lam', 'e', 'lam@cs.stanford.edu'), ('lam', 'p', '650-725-3714'), ('lam', 'p', '650-725-6949'), ('latombe', 'e', 'asandra@cs.stanford.edu'), ('latombe', 'e', 'latombe@cs.stanford.edu'), ('latombe', 'e', 'liliana@cs.stanford.edu'), ('latombe', 'p', '650-721-6625'), ('latombe', 'p', '650-723-0350'), ('latombe', 'p', '650-723-4137'), ('latombe', 'p', '650-725-1449'), ('levoy', 'e', 'ada@graphics.stanford.edu'), ('levoy', 'e', 'melissa@graphics.stanford.edu'), ('levoy', 'p', '650-723-0833'), ('levoy', 'p', '650-724-6865'), ('levoy', 'p', '650-725-3724'), ('levoy', 'p', '650-725-4089'), ('manning', 'e', 'dbarros@cs.stanford.edu'), ('manning', 'e', 'manning@cs.stanford.edu'), ('manning', 'p', '650-723-7683'), ('manning', 'p', '650-725-1449'), ('manning', 'p', '650-725-3358'), ('nass', 'e', 'nass@stanford.edu'), ('nass', 'p', '650-723-5499'), ('nass', 'p', '650-725-2472'), ('nick', 'e', 'nick.parlante@cs.stanford.edu'), ('nick', 'p', '650-725-4727'), ('ouster', 'e', 'ouster@cs.stanford.edu'), ('ouster', 'e', 'teresa.lynn@stanford.edu'), ('pal', 'e', 'pal@cs.stanford.edu'), ('psyoung', 'e', 'patrick.young@stanford.edu'), ('rajeve', 'p', '650-723-4671'), ('rajeve', 'p', '650-723-4648'), ('rajeve', 'p', '650-725-4671'), ('rinard', 'e', 'rinard@lcs.mit.edu'), ('rinard', 'p', '657-250-1221'), ('rinard', 'p', '657-250-6922'), ('serafim', 'e', 'serafim@cs.stanford.edu'), ('serafim', 'p', '650-723-3334'), ('serafim', 'p', '650-725-1449'), ('shoham', 'e', 'shoham@stanford.edu'), ('shoham', 'p', '650-723-3632'), ('shoham', 'p', '650-725-1449'), ('subh', 'e', 'subh@stanford.edu'), ('subh', 'e', 'uma@cs.stanford.edu'), ('subh', 'p', '650-724-1915'), ('subh', 'p', '650-725-3724'), ('subh', 'p', '650-725-6949'), ('thm', 'e', 'pkrokel@stanford.edu'), ('thm', 'p', '650-725-3283'), ('thm', 'p', '650-725-3636'), ('thm', 'p', '650-725-3938'), ('tim', 'p', '650-725-0147'), ('tim', 'p', '650-725-2348'), ('ullman', 'e', 'support@gradience.com'), ('ullman', 'e', 'ullman@cs.stanford.edu'), ('ullman', 'p', '650-494-8816'), ('ullman', 'p', '650-723-2588'), ('ullman', 'p', '650-725-6882'), ('vladien', 'e', 'vladien@stanford.edu'), ('widom', 'e', 'siroker@cs.stanford.edu'), ('widom', 'e', 'widom@cs.stanford.edu'), ('widom', 'p', '650-723-0872'), ('widom', 'p', '650-723-7690'), ('zelenski', 'p', '650-725-2588'), ('zelenski', 'e', 'zelenski@cs.stanford.edu'), ('zelenski', 'p', '650-723-6892'), ('zelenski', 'p', '650-725-8556'), ('zlm', 'e', 'mann@cs.stanford.edu'), ('zlm', 'p', '650-723-4364'), ('zlm', 'p', '650-725-4671')} | EDU in uppercase | {('cheriton', 'e', 'uma@cs.stanford.edu'), ('dabo', 'e', 'dabo@cs.stanford.edu')} | Yes | | |
| {('latombe', 'p', '650-721-6625'), ('latombe', 'p', '650-723-0350'), ('latombe', 'p', '650-723-4137'), ('latombe', 'p', '650-725-1449'), ('levoy', 'e', 'ada@graphics.stanford.edu'), ('levoy', 'e', 'melissa@graphics.stanford.edu'), ('levoy', 'p', '650-723-0833'), ('levoy', 'p', '650-724-6865'), ('levoy', 'p', '650-725-3724'), ('levoy', 'p', '650-725-4089'), ('manning', 'e', 'dbarros@cs.stanford.edu'), ('manning', 'e', 'manning@cs.stanford.edu'), ('manning', 'p', '650-723-7683'), ('manning', 'p', '650-725-1449'), ('manning', 'p', '650-725-3358'), ('nass', 'e', 'nass@stanford.edu'), ('nass', 'p', '650-723-5499'), ('nass', 'p', '650-725-2472'), ('nick', 'e', 'nick.parlante@cs.stanford.edu'), ('nick', 'p', '650-725-4727'), ('ouster', 'e', 'ouster@cs.stanford.edu'), ('ouster', 'e', 'teresa.lynn@stanford.edu'), ('pal', 'e', 'pal@cs.stanford.edu'), ('psyoung', 'e', 'patrick.young@stanford.edu'), ('rajeve', 'p', '650-723-4671'), ('rajeve', 'p', '650-723-4648'), ('rajeve', 'p', '650-725-4671'), ('rinard', 'e', 'rinard@lcs.mit.edu'), ('rinard', 'p', '657-250-1221'), ('rinard', 'p', '657-250-6922'), ('serafim', 'e', 'serafim@cs.stanford.edu'), ('serafim', 'p', '650-723-3334'), ('serafim', 'p', '650-725-1449'), ('shoham', 'e', 'shoham@stanford.edu'), ('shoham', 'p', '650-723-3632'), ('shoham', 'p', '650-725-1449'), ('subh', 'e', 'subh@stanford.edu'), ('subh', 'e', 'uma@cs.stanford.edu'), ('subh', 'p', '650-724-1915'), ('subh', 'p', '650-725-3724'), ('subh', 'p', '650-725-6949'), ('thm', 'e', 'pkrokel@stanford.edu'), ('thm', 'p', '650-725-3283'), ('thm', 'p', '650-725-3636'), ('thm', 'p', '650-725-3938'), ('tim', 'p', '650-725-0147'), ('tim', 'p', '650-725-2348'), ('ullman', 'e', 'support@gradience.com'), ('ullman', 'e', 'ullman@cs.stanford.edu'), ('ullman', 'p', '650-494-8816'), ('ullman', 'p', '650-723-2588'), ('ullman', 'p', '650-725-6882'), ('vladien', 'e', 'vladien@stanford.edu'), ('widom', 'e', 'siroker@cs.stanford.edu'), ('widom', 'e', 'widom@cs.stanford.edu'), ('widom', 'p', '650-723-0872'), ('widom', 'p', '650-723-7690'), ('zelenski', 'p', '650-725-2588'), ('zelenski', 'e', 'zelenski@cs.stanford.edu'), ('zelenski', 'p', '650-723-6892'), ('zelenski', 'p', '650-725-8556'), ('zlm', 'e', 'mann@cs.stanford.edu'), ('zlm', 'p', '650-723-4364'), ('zlm', 'p', '650-725-4671')} | EDU in lowercase | {('cheriton', 'e', 'uma@cs.stanford.edu'), ('dabo', 'e', 'dabo@cs.stanford.edu')} | Yes | | |
| {('latombe', 'p', '650-721-6625'), ('latombe', 'p', '650-723-0350'), ('latombe', 'p', '650-723-4137'), ('latombe', 'p', '650-725-1449'), ('levoy', 'e', 'ada@graphics.stanford.edu'), ('levoy', 'e', 'melissa@graphics.stanford.edu'), ('levoy', 'p', '650-723-0833'), ('levoy', 'p', '650-724-6865'), ('levoy', 'p', '650-725-3724'), ('levoy', 'p', '650-725-4089'), ('manning', 'e', 'dbarros@cs.stanford.edu'), ('manning', 'e', 'manning@cs.stanford.edu'), ('manning', 'p', '650-723-7683'), ('manning', 'p', '650-725-1449'), ('manning', 'p', '650-725-3358'), ('nass', 'e', 'nass@stanford.edu'), ('nass', 'p', '650-723-5499'), ('nass', 'p', '650-725-2472'), ('nick', 'e', 'nick.parlante@cs.stanford.edu'), ('nick', 'p', '650-725-4727'), ('ouster', 'e', 'ouster@cs.stanford.edu'), ('ouster', 'e', 'teresa.lynn@stanford.edu'), ('pal', 'e', 'pal@cs.stanford.edu'), ('psyoung', 'e', 'patrick.young@stanford.edu'), ('rajeve', 'p', '650-723-4671'), ('rajeve', 'p', '650-723-4648'), ('rajeve', 'p', '650-725-4671'), ('rinard', 'e', 'rinard@lcs.mit.edu'), ('rinard', 'p', '657-250-1221'), ('rinard', 'p', '657-250-6922'), ('serafim', 'e', 'serafim@cs.stanford.edu'), ('serafim', 'p', '650-723-3334'), ('serafim', 'p', '650-725-1449'), ('shoham', 'e', 'shoham@stanford.edu'), ('shoham', 'p', '650-723-3632'), ('shoham', 'p', '650-725-1449'), ('subh', 'e', 'subh@stanford.edu'), ('subh', 'e', 'uma@cs.stanford.edu'), ('subh', 'p', '650-724-1915'), ('subh', 'p', '650-725-3724'), ('subh', 'p', '650-725-6949'), ('thm', 'e', 'pkrokel@stanford.edu'), ('thm', 'p', '650-725-3283'), ('thm', 'p', '650-725-3636'), ('thm', 'p', '650-725-3938'), ('tim', 'p', '650-725-0147'), ('tim', 'p', '650-725-2348'), ('ullman', 'e', 'support@gradience.com'), ('ullman', 'e', 'ullman@cs.stanford.edu'), ('ullman', 'p', '650-494-8816'), ('ullman', 'p', '650-723-2588'), ('ullman', 'p', '650-725-6882'), ('vladien', 'e', 'vladien@stanford.edu'), ('widom', 'e', 'siroker@cs.stanford.edu'), ('widom', 'e', 'widom@cs.stanford.edu'), ('widom', 'p', '650-723-0872'), ('widom', 'p', '650-723-7690'), ('zelenski', 'p', '650-725-2588'), ('zelenski', 'e', 'zelenski@cs.stanford.edu'), ('zelenski', 'p', '650-723-6892'), ('zelenski', 'p', '650-725-8556'), ('zlm', 'e', 'mann@cs.stanford.edu'), ('zlm', 'p', '650-723-4364'), ('zlm', 'p', '650-725-4671')} | Space before and after @ | {('dabo', 'e', 'dabo@cs.stanford.edu')} | Yes | | |