

```

import re
import numpy as np
import string
import pandas as pd
import matplotlib as mpl
import matplotlib.pyplot as plt
%matplotlib inline

from subprocess import check_output
from wordcloud import WordCloud, STOPWORDS

stopwords = set(STOPWORDS)
sentences = """We are about to study the idea of a computational process.
Computational processes are abstract beings that inhabit computers.
As they evolve, processes manipulate other abstract things called data.
The evolution of a process is directed by a pattern of rules
called a program. People create programs to direct processes. In effect,
we conjure the spirits of the computer with our spells."""

wordcloud = WordCloud(
    background_color='white',
    stopwords=stopwords,
    max_words=200,
    max_font_size=40,
    random_state=42
).generate(sentences)

fig, axes = plt.subplots(nrows=1, ncols=1, figsize=(10, 10))
axes.imshow(wordcloud)
axes.axis('off')
fig.tight_layout()

```



```

sentences = """We are about to study the idea of a computational process.
Computational processes are abstract beings that inhabit computers.
As they evolve, processes manipulate other abstract things called data.
The evolution of a process is directed by a pattern of rules

```

called a program. People create programs to direct processes. In effect, we conjure the spirits of the computer with our spells."""

```
# remove special characters
```

```
sentences = re.sub('[^A-Za-z0-9]+', ' ', sentences)
```

```
# remove 1 letter words
```

```
sentences = re.sub(r'(?<^| )\w(?:$| )', ' ', sentences).strip()
```

```
# lower all characters
```

```
sentences = sentences.lower()
```

```
print(sentences)
```

we are about to study the idea of computational process computational processes are abstract beings that inhabit computers as they evolve processes manipulate other abstract things called data the evolution of process is directed by pattern of rules called program people create programs to direct processes in effect we conjure the spirits of the computer with our spells

```
words = sentences.split()
```

```
vocab = set(words)
```

```
print(words)
```

```
print(vocab)
```

```
['we', 'are', 'about', 'to', 'study', 'the', 'idea', 'of', 'computational', 'process', 'computational', 'processes', 'are', 'abstract', 'beings', 'that', 'inhabit', 'computers', 'as', 'they', 'evolve', 'processes', 'manipulate', 'other', 'abstract', 'things', 'called', 'data', 'the', 'evolution', 'of', 'process', 'is', 'directed', 'by', 'pattern', 'of', 'rules', 'called', 'program', 'people', 'create', 'programs', 'to', 'direct', 'processes', 'in', 'effect', 'we', 'conjure', 'the', 'spirits', 'of', 'the', 'computer', 'with', 'our', 'spells']
```

```
{'direct', 'they', 'called', 'rules', 'in', 'effect', 'computer', 'computational', 'about', 'manipulate', 'things', 'pattern', 'directed', 'are', 'other', 'people', 'as', 'inhabit', 'evolve', 'create', 'that', 'study', 'the', 'process', 'abstract', 'of', 'idea', 'our', 'data', 'with', 'program', 'programs', 'is', 'by', 'to', 'computers', 'conjure', 'spirits', 'processes', 'evolution', 'we', 'spells', 'beings'}
```

```
vocab_size = len(vocab)
```

```
embed_dim = 10
```

```
context_size = 2
```

```
word_to_ix = {word: i for i, word in enumerate(vocab)}
```

```
ix_to_word = {i: word for i, word in enumerate(vocab)}
```

```
print(word_to_ix)
```

```
print(ix_to_word)
```

```
{'direct': 0, 'they': 1, 'called': 2, 'rules': 3, 'in': 4, 'effect': 5, 'computer': 6, 'computational': 7, 'about': 8, 'manipulate': 9, 'things': 10, 'pattern': 11, 'directed': 12, 'are': 13, 'other': 14, 'people': 15, 'as': 16, 'inhabit': 17, 'evolve': 18, 'create': 19, 'that': 20, 'study': 21, 'the': 22, 'process': 23, 'abstract': 24, 'of': 25, 'idea': 26, 'our': 27, 'data': 28, 'with': 29, 'program': 30, 'programs': 31, 'is': 32, 'by': 33, 'to': 34, 'computers': 35, 'conjure': 36, 'spirits': 37, 'processes': 38, 'evolution': 39, 'we': 40, 'spells': 41, 'beings': 42}
{0: 'direct', 1: 'they', 2: 'called', 3: 'rules', 4: 'in', 5: 'effect', 6: 'computer', 7: 'computational', 8: 'about', 9: 'manipulate', 10: 'things', 11: 'pattern', 12: 'directed', 13: 'are', 14: 'other', 15: 'people', 16: 'as', 17: 'inhabit', 18: 'evolve', 19: 'create', 20: 'that', 21: 'study', 22:
```

```
'the', 23: 'process', 24: 'abstract', 25: 'of', 26: 'idea', 27: 'our', 28:
'data', 29: 'with', 30: 'program', 31: 'programs', 32: 'is', 33: 'by', 34:
'to', 35: 'computers', 36: 'conjure', 37: 'spirits', 38: 'processes', 39:
'evolution', 40: 'we', 41: 'spells', 42: 'beings'}
```

```
# data - [(context), target]
```

```
data = []
```

```
for i in range(2, len(words) - 2):
```

```
    context = [words[i - 2], words[i - 1], words[i + 1], words[i + 2]]
```

```
    target = words[i]
```

```
    data.append((context, target))
```

```
print(data[:5])
```

```
((['we', 'are', 'to', 'study'], 'about'), (['are', 'about', 'study', 'the'],
'to'), (['about', 'to', 'the', 'idea'], 'study'), (['to', 'study', 'idea',
'of'], 'the'), (['study', 'the', 'of', 'computational'], 'idea'))
```

```
embeddings = np.random.random_sample((vocab_size, embed_dim))
```

```
print(embeddings)
```

```
[[0.55902317 0.8796336 0.01606508 0.83640464 0.94515038 0.03671365
 0.27138932 0.69865781 0.7621218 0.49927705]
 [0.10055462 0.63714277 0.42372804 0.478108 0.26711547 0.09149576
 0.29168269 0.94036423 0.22413572 0.39162099]
 [0.95464716 0.38335215 0.68628304 0.01724722 0.65466762 0.82630526
 0.14338801 0.47979158 0.55403334 0.59806094]
 [0.84383432 0.23158652 0.49851811 0.52690493 0.19314359 0.03769864
 0.9239182 0.58604157 0.29585361 0.75399642]
 [0.30806521 0.34873435 0.68050268 0.88605163 0.34890095 0.04797554
 0.69039153 0.86210584 0.19452842 0.27652878]
 [0.44921168 0.01018458 0.10001453 0.36017239 0.37986563 0.57077681
 0.36082491 0.15131566 0.79828081 0.21166546]
 [0.10538946 0.86263733 0.97936195 0.80499425 0.47077666 0.15855748
 0.20198191 0.74666776 0.11708653 0.26319007]
 [0.96836749 0.51423675 0.46540649 0.4272708 0.45180645 0.78393261
 0.93397244 0.53718584 0.36399799 0.12210278]
 [0.29130761 0.14499853 0.85154888 0.44465339 0.58324302 0.64895118
 0.13842134 0.52480933 0.69221816 0.15314234]
 [0.89677092 0.08651211 0.63968209 0.17938769 0.92941206 0.96729667
 0.93935048 0.90278976 0.61506383 0.55344748]
 [0.47560007 0.57787429 0.50357385 0.99180738 0.03867795 0.61980775
 0.42743254 0.30866664 0.57509802 0.72280169]
 [0.0292131 0.18008654 0.75532664 0.18913698 0.48467134 0.49147614
 0.99957892 0.1025795 0.31481234 0.59640461]
 [0.34053374 0.88620595 0.32149636 0.75317574 0.90496078 0.16534758
 0.11691434 0.5436065 0.42417629 0.16822331]
 [0.63267794 0.82218887 0.50497123 0.57612382 0.26080854 0.08491206
 0.92220003 0.99517395 0.65428914 0.56916061]
 [0.59867719 0.32500871 0.01006457 0.46699615 0.91691341 0.44061327
 0.69203526 0.19498461 0.27813834 0.48776731]
 [0.03737713 0.76213107 0.29010761 0.30563751 0.68342246 0.64852369
 0.3830537 0.62640629 0.54361764 0.99560582]
 [0.71125134 0.32188125 0.19249123 0.5940112 0.65361697 0.5670915
 0.06405644 0.89447002 0.02501037 0.02569859]
 [0.70682621 0.1697142 0.45092039 0.88107085 0.61191993 0.39235914
 0.92079698 0.62779889 0.16279128 0.81865555]
 [0.5687993 0.22804503 0.93789164 0.07070246 0.35066001 0.53565905
 0.64651151 0.61669068 0.54499238 0.7218042 ]]
```

[0.39141204 0.95963752 0.08899486 0.57230914 0.15191606 0.07459051
0.16198143 0.34555402 0.8758638 0.28783637]
[0.61501984 0.37891144 0.38973929 0.68123779 0.66492461 0.29816377
0.9266779 0.56303613 0.53299404 0.06964219]
[0.06740883 0.41160211 0.50818432 0.57372577 0.66085666 0.1742478
0.20957445 0.52229692 0.60760105 0.05414961]
[0.35368909 0.35298248 0.96208777 0.18958784 0.65711821 0.33987877
0.26141294 0.03329625 0.63378331 0.34383058]
[0.55568049 0.19620255 0.09678858 0.88879468 0.00218402 0.85234013
0.55104576 0.82430225 0.8379663 0.93048525]
[0.92481104 0.98222132 0.80940994 0.2548421 0.64036237 0.38058139
0.95220029 0.3352988 0.54985922 0.64149738]
[0.49328692 0.50287429 0.03606609 0.88675683 0.53395749 0.45546944
0.16167272 0.79382407 0.06654294 0.82510425]
[0.72163701 0.2467374 0.7303794 0.25640706 0.12220016 0.9165005
0.10952576 0.91633619 0.96842452 0.73890695]
[0.36586411 0.7515007 0.14722247 0.6406858 0.68766514 0.21646427
0.14877631 0.72782687 0.14050458 0.74958574]
[0.13573147 0.91991145 0.13750531 0.88963138 0.34709709 0.76010183
0.09644531 0.51585206 0.57848192 0.3682413]
[0.35574132 0.06437098 0.01449723 0.01749394 0.99009996 0.87998187
0.47205441 0.85277301 0.79077788 0.0543162]
[0.61285547 0.71782534 0.25427413 0.09433277 0.45767025 0.03140795
0.47322179 0.17184041 0.88348221 0.35366897]
[0.69421424 0.23188181 0.76184257 0.61563732 0.54398131 0.25470129
0.57724233 0.52300325 0.05717392 0.63644609]
[0.52694406 0.4356314 0.45366411 0.09911894 0.59120233 0.37231263
0.25315644 0.36820299 0.01271292 0.70130125]
[0.89182372 0.98752574 0.81202508 0.64138494 0.05069737 0.21197388
0.38676834 0.55902303 0.53836614 0.60716916]
[0.60786441 0.92728581 0.69198122 0.61599739 0.74123805 0.35148439
0.37519341 0.62511948 0.5415566 0.90589371]
[0.16547332 0.98042449 0.78742008 0.17028886 0.54396077 0.14550411
0.94306519 0.41288371 0.30342499 0.39803867]
[0.46223938 0.29348248 0.04374014 0.40268991 0.06706729 0.39876046
0.17708012 0.29227249 0.3954279 0.67426925]
[0.91588121 0.08722407 0.48936658 0.96222817 0.25489614 0.23560044
0.39150502 0.1000599 0.72920947 0.3170044]
[0.37038171 0.10456142 0.9623594 0.44572792 0.31929118 0.95818652
0.70486655 0.28360931 0.87294152 0.95132912]
[0.11802475 0.14982839 0.56022318 0.40463393 0.06156818 0.44766894
0.63876637 0.35104739 0.24342893 0.51628805]
[0.36886228 0.97610319 0.9001242 0.07279782 0.31235222 0.75657946
0.20040641 0.12424929 0.56093482 0.53550702]
[0.21426268 0.89639685 0.45833907 0.01409733 0.07841329 0.51662104
0.30730498 0.29657929 0.86601284 0.71593723]
[0.43256503 0.82074468 0.45736564 0.95366652 0.88362831 0.75057239
0.84926887 0.02484717 0.56286909 0.85824221]]