Natural Language Processing

with NuPIC

NLTK

Python Natural Language Tool Kit

NLTK Part of Speech Tagging

That	DET	determiner
is	V	verb
the	DET	determiner
way	N	noun
Thor	NP	proper noun
got	VD	past tense
his	PRO	pronoun
wonderful	ADJ	adjective
hammer	N	noun
.	.	.

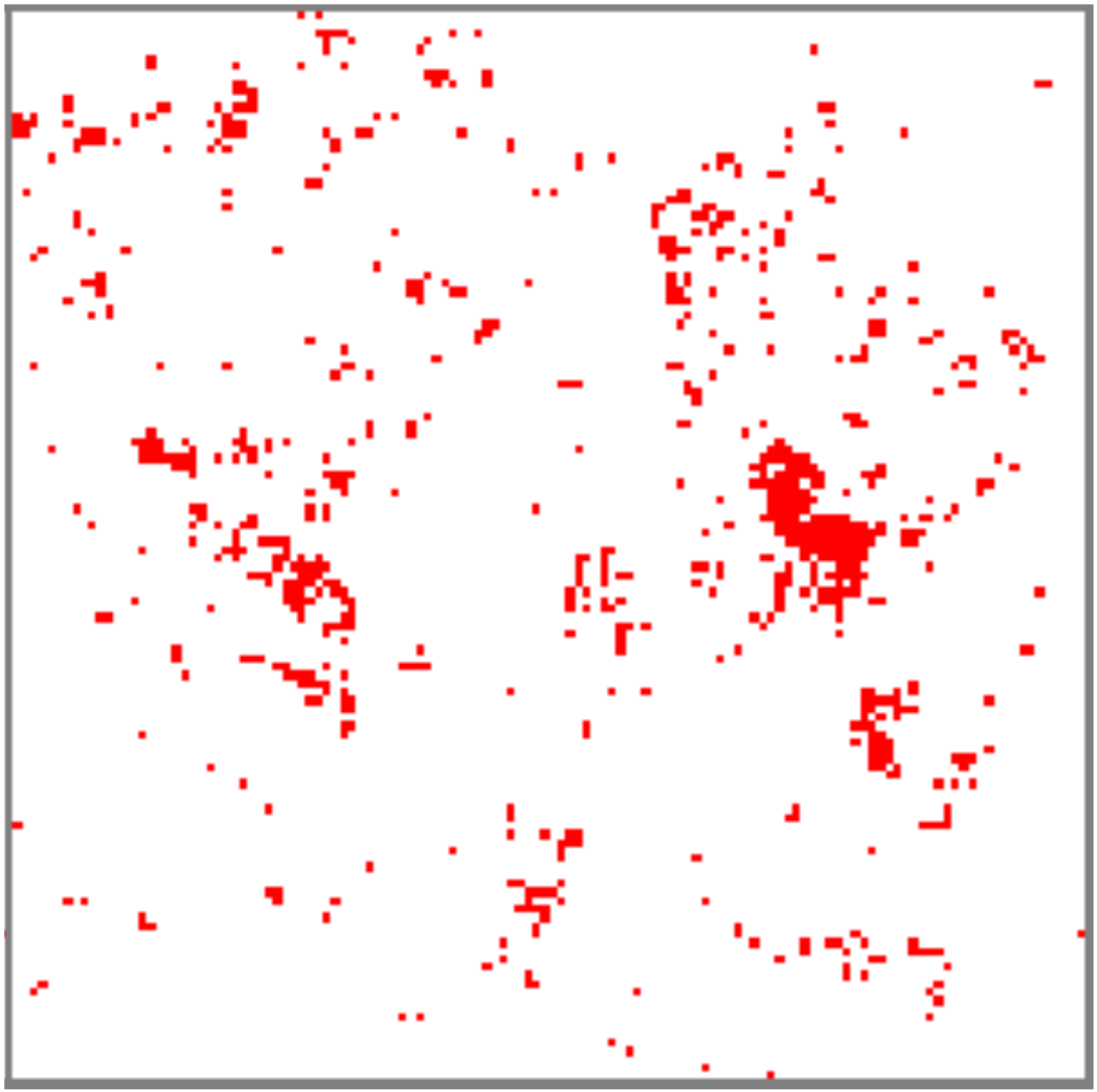
nupic_nlp

word POS demo

CEPT

Word SDRs

“metal”



“metal”

{'**height**': 128, '**width**': 128, '**sparsity**': 5.0, '**positions**': [34, 36, 292, 293, 294, 295, 308, 311, 421, 424, 433, 549, 560, 607, 656, 784, 794, 802, 807, 945, 946, 947, 952, 1050, 1051, 1074, 1076, 1080, 1146, 1147, 1178, 1179, 1180, 1286, 1308, 1414, 1425, 1426, 1432, 1434, 1435, 1436, 1504, 1505, 1536, 1538, 1550, 1552, 1561, 1579, 1581, **1664 … 12524**, 12525, 12526, 12527, 12603, 12607, 12610, 12611, 12737, 12738, 12739, 12852, 12865, 12902, 12993, 13009, 13098, 13371, 13372, 13377, 13470, 13471, 13501, 13502, 13503, 13504, 13574, 13576, 13599, 13606, 13629, 13633, 13650, 13756, 13757, 13758, 13759, 13839, 13861, 13887, 13967, 13968, 14014, 14038, 14142, 14166, 14180, 14207, 14266, 14296, 14302, 14305, 14306, 14309, 14315, 14430, 14435, 14443, 14444, 14445, 14446, 14522, 14555, 14566, 14567, 14648, 14691, 14703, 14781, 14819, 14821, 14851, 14909, 14910, 14958, 14978, 15050, 15085, 15214, 15406, 15408, 15469, 15815, 15945, 16210, 16346]}

“metal”

“metal”

{'height': 128, 'width': 128, 'sparsity': 5.0,
'positions': [34, 36, 292, 293, 294, 295,
308, 311, 421, 424, 433, 549, 560, 607,
656, 784, 794, 802, 807, 945, 946, 947,
952, 1050, 1051, 1074, 1076, 1080, 1146,
1147, 1178, 1179, 1180, 1286, 1308,
1414, 1581, 1664 ... 13629, 13633,
13650, 13756, 14306, 14309, 14315,
14430, 14435, 14443, 14444, 14445,
14446, 14522, 14555, 14566, 14567,
14648, 14691, 14703, 14781, 14819,
14821, 14851, 14909, 14910, 14958,
14978, 15050, 15085, 15214, 15406,
15408, 15469, 15815, 15945, 16210,
16346]}

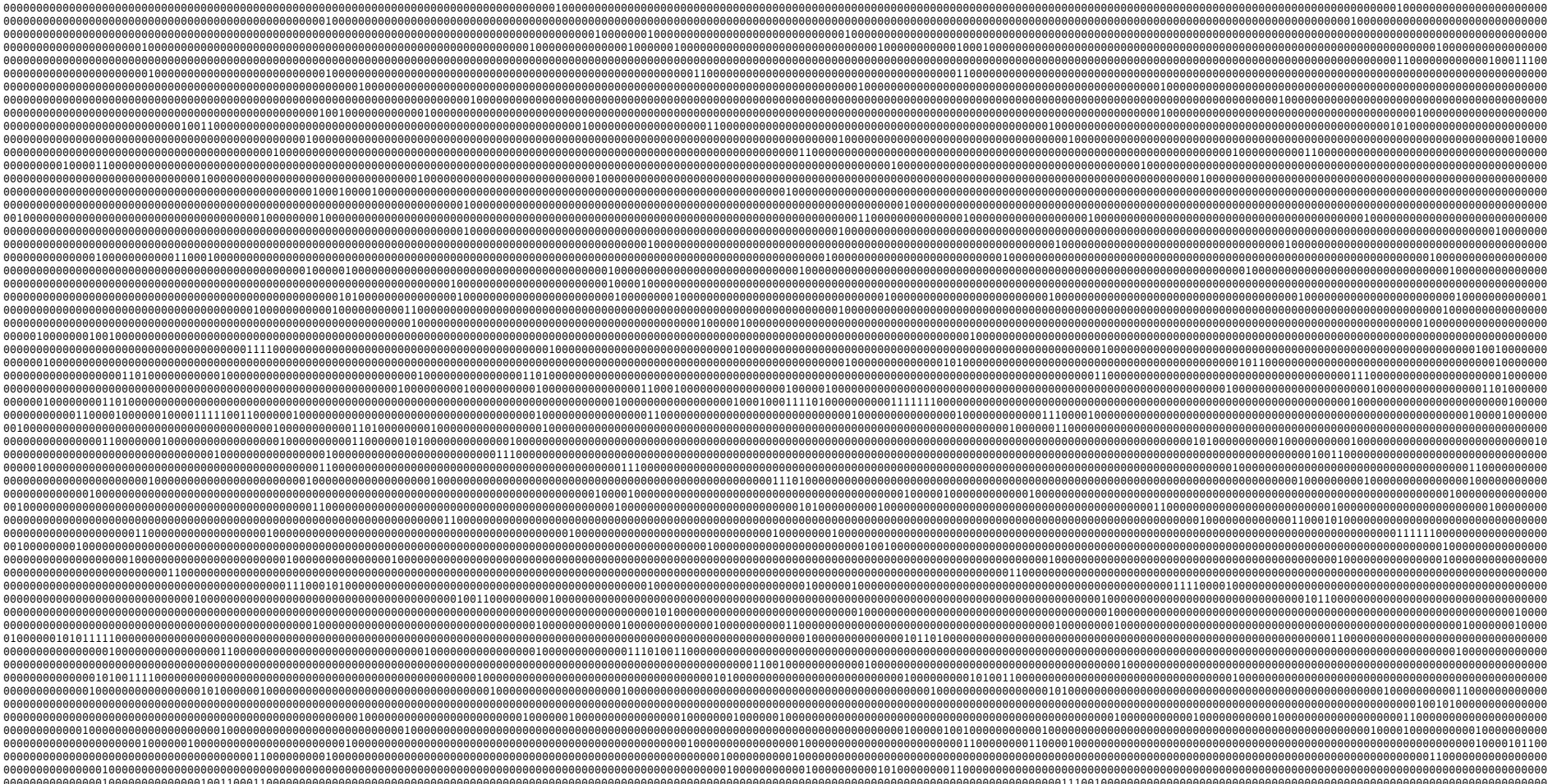
metal | 0
metals | 1
materials | 2
metallic | 3
material | 4
alloys | 5
solid | 6
oxide | 7
aluminium | 8
copper | 9
corrosion | 10
oxides | 11
thin | 12
steel | 13
alloy | 14
chemical | 15
nickel | 16
platinum | 17
carbon | 18
halides | 19
heavy metal | 20
oxidation | 21
plastic | 22
heating | 23

pycept

Python CEPT API Client

pycept: get SDR from term

```
import pycept
ceptClient = pycept.Cept("your_app_id", "your_app_key")
catSdr = ceptClient.getSdr("cat")
```



pycept: get bitmap from term

```
catBitmap = ceptClient.getBitmap ("cat")
```


{u'height': 128, u'width': 128, u'sparsity': 4.0, u'positions': [84, 212, 284, 440, 560, 568, 598, 726, 785, 800,
807, 838, 850, 854, 923, 1152, 1153, 1166, 1170, 1171, 1172, 1197, 1224, 1280, 1281, 1320, 1321, 1464, 1540, 1586,
1716, 1839, 1928, 1931, 1944, 2056, 2095, 2142, 2145, 2146, 2203, 2222, 2223, 2274, 2326, 2328, 2396, 2477, 2512,
2580, 2626, 2706, 2707, 2772, 2783, 2784, 2829, 2834, 2835, 2955, 2993, 3085, 3118, 3145, 3237, 3337, 3341, 3346,
3409, 3595, 3662, 3762, 3799, 3808, 3890, 3891, 3906, 3925, 3967, 4065, 4122, 4222, 4328, 4390, 4425, 4479, 4491,
4492, 4496, 4590, 4617, 4682, 4746, 4752, 4792, 4821, 4889, 4920, 5003, 5027, 5032, 5221, 5223, 5239, 5263, 5272,
5304, 5329, 5367, 5391, 5404, 5443, 5455, 5466, 5467, 5532, 5624, 5702, 5746, 5752, 5856, 5880, 5888, 5891, 6022,
6147, 6148, 6149, 6150, 6193, 6221, 6277, 6334, 6337, 6351, 6473, 6488, 6490, 6533, 6535, 6536, 6572, 6598, 6599,
6601, 6613, 6643, 6659, 6660, 6662, 6746, 6747, 6785, 6786, 6787, 6808, 6875, 6885, 6896, 6912, 6913, 6917, 6934,
6940, 7001, 7023, 7040, 7041, 7043, 7056, 7065, 7066, 7068, 7143, 7161, 7165, 7169, 7170, 7171, 7172, 7174, 7185,
7186, 7187, 7188, 7189, 7190, 7191, 7255, 7279, 7296, 7297, 7302, 7309, 7314, 7315, 7316, 7317, 7318, 7321, 7322,
7329, 7366, 7383, 7384, 7414, 7430, 7443, 7444, 7445, 7450, 7508, 7513, 7522, 7561, 7573, 7574, 7576, 7585, 7602,
7673, 7680, 7681, 7770, 7771, 7779, 7797, 7808, 7809, 7816, 7818, 7832, 7936, 7938, 7949, 7960, 7988, 8022, 8039,
8065, 8066, 8067, 8189, 8192, 8193, 8230, 8273, 8274, 8319, 8320, 8321, 8412, 8448, 8449, 8482, 8506, 8525, 8577,
8578, 8579, 8581, 8657, 8667, 8683, 8708, 8785, 8790, 8832, 8838, 8851, 8915, 8932, 8977, 8978, 9023, 9051, 9053,
9063, 9105, 9106, 9132, 9156, 9232, 9233, 9347, 9361, 9362, 9366, 9368, 9420, 9421, 9440, 9486, 9517, 9526, 9612,
9613, 9614, 9615, 9616, 9617, 9637, 9646, 9742, 9766, 9769, 9854, 9889, 9913, 9929, 10029, 10073, 10089, 10130,
10131, 10258, 10259, 10383, 10384, 10385, 10389, 10391, 10438, 10462, 10469, 10518, 10519, 10520, 10521, 10526,
10604, 10644, 10647, 10648, 10658, 10742, 10773, 10775, 10776, 10909, 10911, 10940, 10978, 11040, 11092, 11126,
11139, 11153, 11164, 11165, 11205, 11214, 11267, 11275, 11281, 11288, 11290, 11292, 11293, 11294, 11295, 11296,
11402, 11417, 11419, 11420, 11422, 11482, 11483, 11531, 11548, 11549, 11579, 11596, 11610, 11611, 11612, 11614,
11617, 11618, 11736, 11864, 11865, 11868, 11881, 11920, 11999, 12001, 12004, 12005, 12006, 12007, 12057, 12093,
12095, 12122, 12132, 12134, 12137, 12138, 12172, 12233, 12250, 12252, 12259, 12294, 12314, 12361, 12379, 12381,
12430, 12441, 12442, 12670, 12673, 12675, 12744, 12769, 12776, 12793, 12801, 12808, 12859, 12871, 12883, 12903,
12904, 12937, 12958, 12986, 13062, 13068, 13071, 13083, 13133, 13138, 13149, 13181, 13188, 13212, 13264, 13281,
13306, 13307, 13316, 13317, 13322, 13384, 13389, 13391, 13392, 13433, 13434, 13444, 13504, 13515, 13612, 13613,
13645, 13740, 13752, 13763, 13765, 13774, 13775, 13880, 13895, 13898, 13899, 13903, 13904, 14026, 14027, 14028,
14031, 14104, 14140, 14149, 14157, 14160, 14198, 14213, 14260, 14266, 14268, 14269, 14287, 14291, 14309, 14311,
14410, 14422, 14512, 14522, 14555, 14563, 14564, 14640, 14648, 14649, 14652, 14678, 14691, 14692, 14693, 14694,
14704, 14775, 14781, 14821, 14822, 14910, 14911, 14931, 15023, 15024, 15068, 15069, 15075, 15086, 15150, 15151,
15179, 15193, 15214, 15307, 15319, 15406, 15415, 15435, 15442, 15505, 15542, 15563, 15571, 15580, 15584, 15641,
15691, 15692, 15696, 15699, 15709, 15710, 15711, 15712, 15714, 15720, 15764, 15766, 15767, 15768, 15769, 15784,

pycept: similar terms from bitmap

```
similar = ceptClient.bitmapToTerms(catBitmap)
```


```
[{'term': u'cat', 'rank': 0}, {'term': u'cats', 'rank': 1}, {'term': u'dog', 'rank': 2}, {'term': u'dogs', 'rank':
3}, {'term': u'animals', 'rank': 4}, {'term': u'animal', 'rank': 5}, {'term': u'rabbits', 'rank': 6}, {'term':
u'pet', 'rank': 7}, {'term': u'feline', 'rank': 8}, {'term': u'rabbit', 'rank': 9}, {'term': u'pets', 'rank': 10},
{'term': u'mouse', 'rank': 11}, {'term': u'canine', 'rank': 12}, {'term': u'mice', 'rank': 13}, {'term': u'mammals',
'rank': 14}, {'term': u'rodents', 'rank': 15}, {'term': u'prey', 'rank': 16}, {'term': u'canines', 'rank': 17},
{'term': u'humans', 'rank': 18}, {'term': u'monkey', 'rank': 19}, {'term': u'adult', 'rank': 20}, {'term': u'other
animals', 'rank': 21}, {'term': u'hares', 'rank': 22}, {'term': u'mammal', 'rank': 23}, {'term': u'predators',
'rank': 24}, {'term': u'eyes', 'rank': 25}, {'term': u'breeds', 'rank': 26}, {'term': u'wild', 'rank': 27}, {'term':
u'carnivores', 'rank': 28}, {'term': u'rodent', 'rank': 29}, {'term': u'birds', 'rank': 30}, {'term': u'fictional',
'rank': 31}, {'term': u'species', 'rank': 32}, {'term': u'skin', 'rank': 33}, {'term': u'animated', 'rank': 34},
{'term': u'parasites', 'rank': 35}, {'term': u'claws', 'rank': 36}, {'term': u'teeth', 'rank': 37}, {'term': u'rare',
'rank': 38}, {'term': u'bites', 'rank': 39}, {'term': u'rats', 'rank': 40}, {'term': u'feral', 'rank': 41}, {'term':
u'like', 'rank': 42}, {'term': u'invasive', 'rank': 43}, {'term': u'monkeys', 'rank': 44}, {'term': u'ears', 'rank':
45}, {'term': u'anthropomorphic', 'rank': 46}, {'term': u'pigs', 'rank': 47}, {'term': u'appear', 'rank': 48},
{'term': u'red list', 'rank': 49}, {'term': u'reptiles', 'rank': 50}, {'term': u'eat', 'rank': 51}, {'term': u'rat',
'rank': 52}, {'term': u'legs', 'rank': 53}, {'term': u'eating', 'rank': 54}, {'term': u'carnivorans', 'rank': 55},
{'term': u'females', 'rank': 56}, {'term': u'tails', 'rank': 57}, {'term': u'scavengers', 'rank': 58}, {'term':
u'breed', 'rank': 59}, {'term': u'hair', 'rank': 60}, {'term': u'fictional characters', 'rank': 61}, {'term':
u'megafauna', 'rank': 62}, {'term': u'breeding', 'rank': 63}, {'term': u'bird', 'rank': 64}, {'term': u'described',
'rank': 65}, {'term': u'characters', 'rank': 66}, {'term': u'fauna', 'rank': 67}, {'term': u'insects', 'rank': 68},
{'term': u'duck', 'rank': 69}, {'term': u'glands', 'rank': 70}, {'term': u'domesticated', 'rank': 71}, {'term':
u'spider', 'rank': 72}, {'term': u'snakes', 'rank': 73}, {'term': u'adults', 'rank': 74}, {'term': u'eggs', 'rank':
75}, {'term': u'lizards', 'rank': 76}, {'term': u'marsupials', 'rank': 77}, {'term': u'biting', 'rank': 78}, {'term':
u'carnivora', 'rank': 79}, {'term': u'predator', 'rank': 80}, {'term': u'raccoon', 'rank': 81}, {'term': u'list',
'rank': 82}, {'term': u'subspecies', 'rank': 83}, {'term': u'nocturnal', 'rank': 84}, {'term': u'other mammals',
'rank': 85}, {'term': u'venomous', 'rank': 86}, {'term': u'bite', 'rank': 87}, {'term': u'foxes', 'rank': 88},
{'term': u'creatures', 'rank': 89}, {'term': u'appearance', 'rank': 90}, {'term': u'genomes', 'rank': 91}, {'term':
u'deer', 'rank': 92}, {'term': u'squirrel', 'rank': 93}, {'term': u'character', 'rank': 94}, {'term': u'tail',
'rank': 95}, {'term': u'frog', 'rank': 96}, {'term': u'creature', 'rank': 97}, {'term': u'carnivorous', 'rank': 98},
{'term': u'bred', 'rank': 99}]
```

pycept: similar terms from bitmap

```
similar = ceptClient.bitmapToTerms(catBitmap)
```

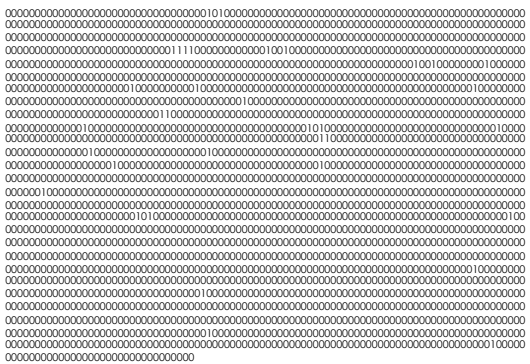
cat | 0
cats | 1
dog | 2
dogs | 3
animals | 4
animal | 5
rabbits | 6
pet | 7
feline | 8
rabbit | 9
pets | 10
mouse | 11
canine | 12
mice | 13
mammals | 14
rodents | 15
prey | 16
canines | 17
humans | 18
monkey | 19
adult | 20
other animals | 21
hares | 22
mammal | 23
predators | 24

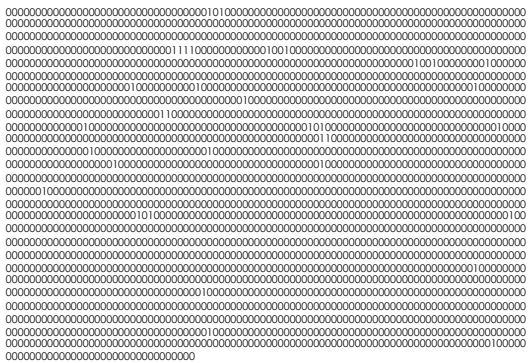
Word association experiment

Two Lists

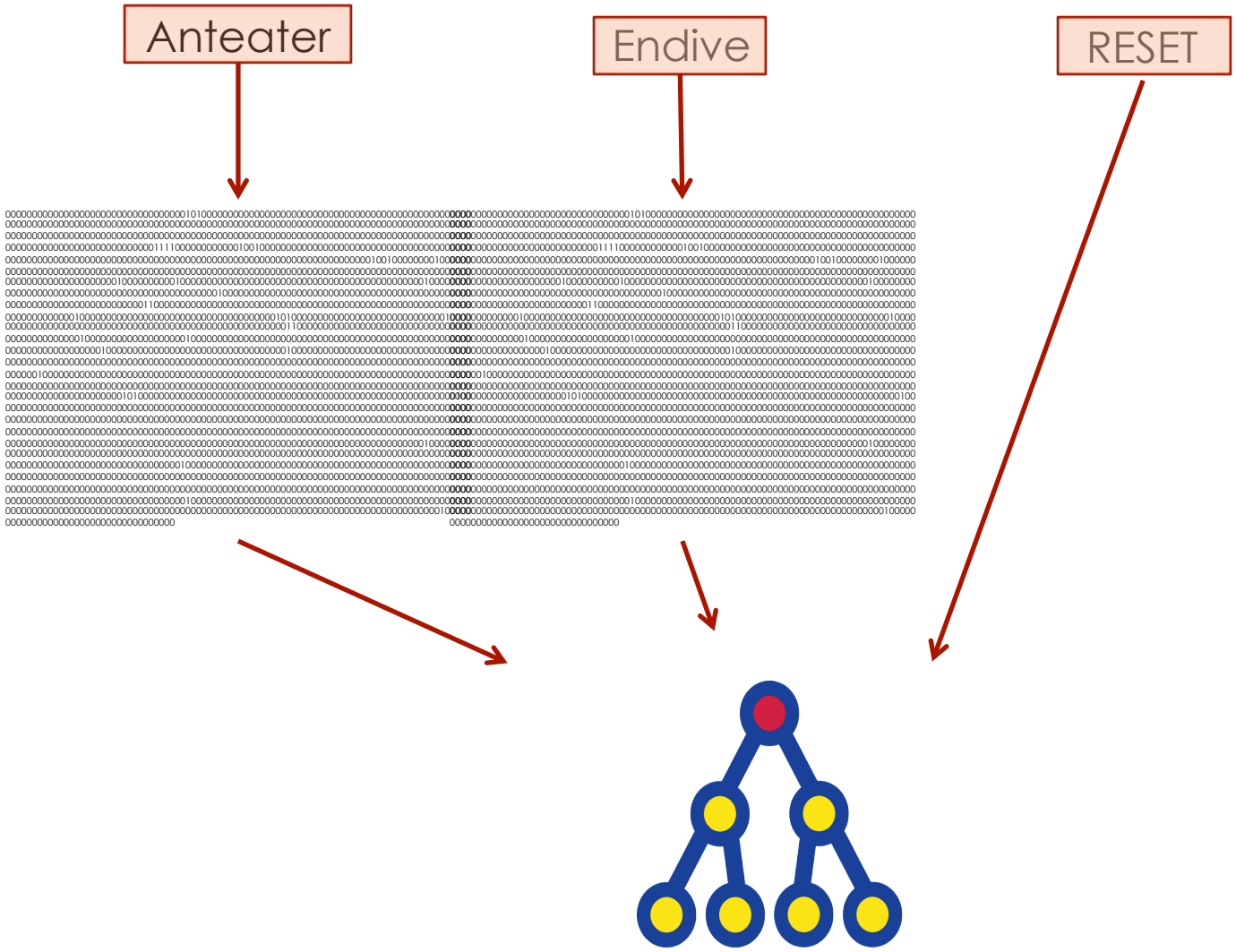
Aardvark
Albatross
Alligator
Alpaca
Ant
Anteater
Antelope
Ape
Armadillo
Donkey
Baboon
Badger
Barracuda
Bat
Bear
Beaver
Bee
Bison
Boar
Buffalo
…

Amaranth
Arugula
Beet
Borage
Cabbage
Catsear
Celery
Celtuce
Chaya
Chickweed
Chicory
Chrysanthemum
Purslane
Corn
Cress
Dandelion
Endive
Fiddlehead
Kale
Komatsuna
…





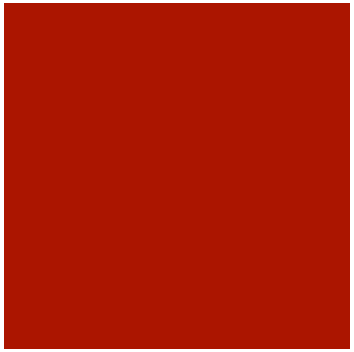
Two Lists



nupic_nlp

word association demo

NLP Resources



- Tools
- Experiments
- Challenges