

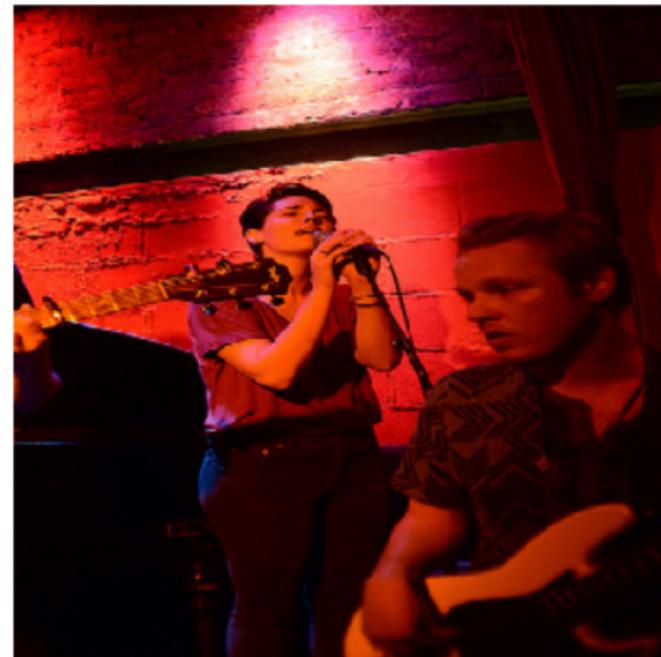
Experiments in image captioning, *With a twist*

The Amusemater Captioner

Joshua Banks Mailman

Experiments in Captioning, *Wit* a Twist

The Amusemater Captioner



Close, but no guitar

Motivation

Caption contest

THE
NEW YORKER

News Books & Culture Fiction & Poetry Humor & Cartoons Magazine Crossword Video Podcasts

"You were constantly nagging me about having some exposed brick in our place!"

"Exposed brick is all the rage."

"You know how you've always wanted exposed brick..."

"Exposed brick, honey! Just increased our property value!"

"Haven't you heard? Exposed brick is out."

"You get the exposed brick you've always wanted, and I get to watch the second half in peace."

"You always said that you wished we had exposed brick."

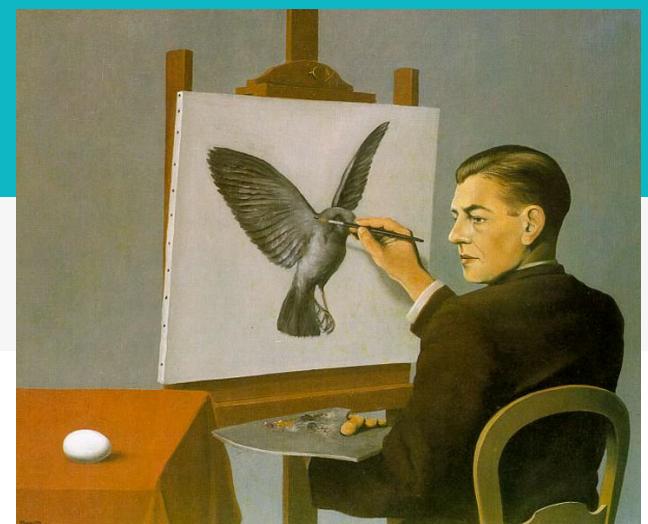
Milton Babbitt's titles

It Takes Twelve to Tango

Whirled Series

The Joy of More Sextets

Rene Magritte's titles



Clairvoyance

Cockney rhyming slang

"Army and navy" (gravy)

As gravy was plentiful at mealtimes in both services.

"Basin of gravy" (baby)

Suggestive of the softness of the foods on which babies are fed.

"Bees and honey" (money)

As bees are the epitome of work, work produces money, the possession of which is sweet.

"Borrow and beg" (egg)

A term that enjoyed a fresh lease of life during the second world war and the food-rationing period.

"Bottle and stopper" (copper)

Policeman. There are a possible pair of inferences: to bottle meaning to enclose and a stopper meaning one who holds another back from a course of action.

"Box of toys" (noise)

As a box of toys, particularly a new one given as a present at Christmas time, causes a great deal of noise to be made.

"Can't keep still" (treadmill)

Since people sentenced to that [19th century punishment](#) could not keep still for a second.

"Clever mike" (bike)

Influenced by the extreme displays that adolescents are inclined to perform on a bicycle, especially when showing off.

"Coals and coke" (broke)

Since both coal and coke used to be supplied in large blocks that had to be broken down before their use.

"Coat and badge" (cadge)

Sometimes known as "Doggett's" as watermen who possessed the Doggett Coat and Badge could charge higher fares than those without.

"Collar and cuff" (puff)

An effeminate man, suggestive either by their neatness of dress, or by "powder puff".

Bringing to mind something not literally pictured or mentioned

Thinking around the corner

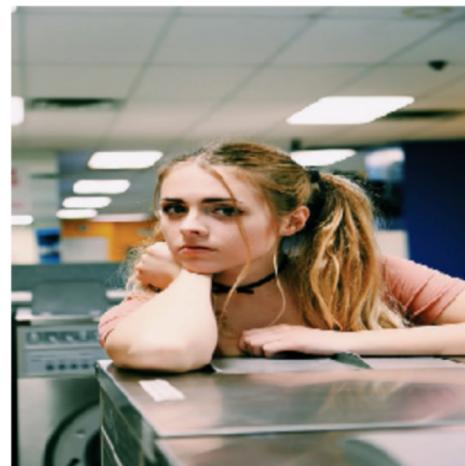
Experiments in Captioning, *Wit* a Twist

The Amusemater Captioner

A horizontal slider bar with a red circular slider. The number '2514' is displayed above the slider. On the left end of the slider, the number '1' is shown. On the right end, the number '5500' is shown.

- Brandish this for broad
 - Touch me for subtlety
 - Click me for contrary craziness

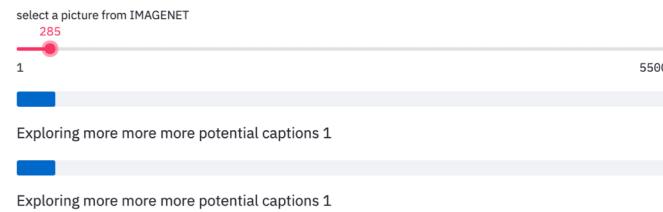
Done!



Smoking gown

Experiments in Captioning, Wit a Twist

The Amusemater Captioner



Let the genie out of the beetle

Show other captions

Experiments in Captioning, *Wit* a Twist

The Amusemater Captioner

select a picture from IMAGENET



Exploring more and more and more potential captions

Exploring more and more and more and more potential captions



Suit yourself in the foot

Data Resources

Data Resources

IMAGENET

5500 with a 1000 labels (categories)

BeautifulSoup

7ESL.com (English as Second Language)

1500 English Idioms

“Always a day late and a dollar short”

“They say a cat has nine lives”

“Giving a hundred and ten percent”

Colloquialize: formal to informal:

“he is” → “he’s”

“do not” → “don’t”

“one will” → “you’ll”

SentiWordNet

>117,000 independent sentences
treated as separate *documents*

Gensim corpuses

Google News

Text 8

CMU (Carnegie Mellon University)

Corpus of English Words

IPA

International Phonetic Alphabet



Screeching to the choir

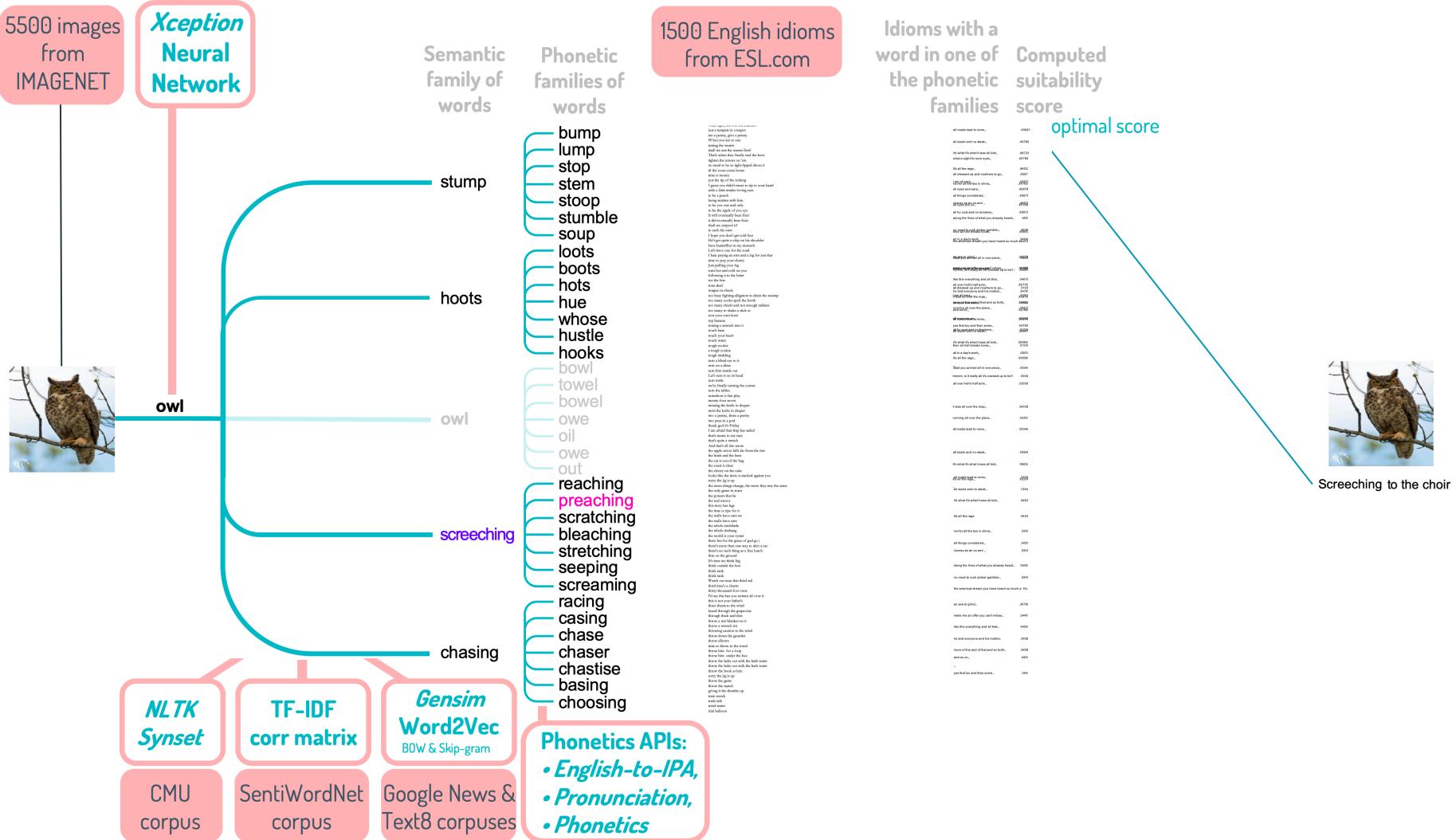


Screeching to the choir



Screeching to the choir

The algorithm and its algorithms

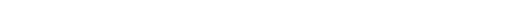


Experiments in Captioning, Wit a Twist

The Amusemater Captioner

Slide to select a picture from IMAGENET

971



1 556

- Brandish this for broad
 - Touch me for subtlety
 - Click me for contrary craziness

Done!



The peel McCoy

Experiments in Captioning, Wit a Twist

The Amusemater Captioner

Slide to select a picture from IMAGENET



- Brandish this for broad
- Touch me for subtlety
- Click me for contrary craziness

Done!



Between the diesel and the deep blue sea

Experiments in Captioning, Wit a Twist

The Amusemater Captioner

Slide to select a picture from IMAGENET



Brandish this for broad

Click me for subtlety

Exploring more and more and more potential captions

Exploring more and more potential captions

Exploring more and more potential captions



Install talk and no trousers



Just a bones throw away

Experiments in image captioning, *With a twist*
The Amusemater Captioner



Just a bones throw away

Joshua Banks Mailman

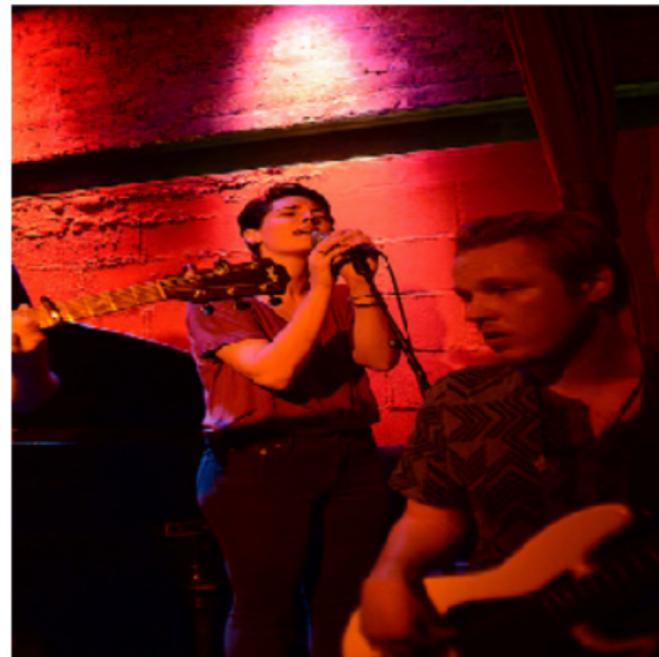
<https://twitter.com/JoshuaBMailman>
joshuabanksmailman@gmail.com

<https://github.com/jmailman>
<https://www.linkedin.com/in/joshua-banks-mailman/>

23

Experiments in Captioning, *Wit a Twist*

The Amusemater Captioner



Close, but no guitar

Experiments in Captioning, *Wit a Twist*

The Amusemater Captioner



Every man for damselfly

Experiments in Captioning, *Wit a Twist*

The Amusemater Captioner



Up, up in the bear



A bear bird

Experiments in Captioning, *Wit a Twist*

The Amusemater Captioner



Pie the knot

Experiments in Captioning, *Wit a Twist*

The Amusemater Captioner



Up, up in the bear

	dist_score	semantic_match	sem_dist	phonetic_match	phon_dist
A bear bird	0.9000	bear	0.6000	rare	0.3000
Back to bear one	0.9000	bear	0.6000	square	0.3000
Put your money bear your mouth is	0.9000	bear	0.6000	where	0.3000
Hightail it (out of bear	0.9000	bear	0.6000	there	0.3000
Turnabout is bear play	0.9000	bear	0.6000	fair	0.3000
Up, up in the bear	0.9000	bear	0.6000	air	0.3000
Bear one's hair out	0.9000	bear	0.6000	tear	0.3000
Hightail it (out of bear	0.9000	bear	0.6000	there	0.3000
Not a bear out of place	0.9000	bear	0.6000	hair	0.3000
If the shoe fits, bear it	0.9000	bear	0.6000	wear	0.3000
Bear the living daylights out of someone	0.9000	bear	0.6000	scare	0.3000
Lion's bear	0.9000	bear	0.6000	share	0.3000
Put your money bear your mouth is	0.9000	bear	0.6000	where	0.3000
Bear the rod and spoil the child	0.9000	bear	0.6000	spare	0.3000
Back to bear one	0.9000	bear	0.6000	square	0.3000
Not have a bear	0.9000	bear	0.6000	prayer	0.3000
Give someone a run for bear money	0.9000	bear	0.6000	their	0.3000
Bear one's hair out	0.9000	bear	0.6000	tear	0.3000
Turnabout is bear play	0.9000	bear	0.6000	fair	0.3000
Up, up in the bear	0.9000	bear	0.6000	air	0.3000
Bear the rod and spoil the child	0.9000	bear	0.6000	spare	0.3000
Lion's bear	0.9000	bear	0.6000	share	0.3000
If the shoe fits, bear it	0.9000	bear	0.6000	wear	0.3000
Bear the living daylights out of someone	0.9000	bear	0.6000	scare	0.3000
Not a bear out of place	0.9000	bear	0.6000	hair	0.3000



Up, up in the bear

	dist_score	semantic_match	sem_dist	phonetic_match	phon_dist
A bear bird	0.9000	bear	0.6000	rare	0.3000
Back to bear one	0.9000	bear	0.6000	square	0.3000
Put your money bear your mouth is	0.9000	bear	0.6000	where	0.3000
Hightail it (out of bear	0.9000	bear	0.6000	there	0.3000
Turnabout is bear play	0.9000	bear	0.6000	fair	0.3000
Up, up in the bear	0.9000	bear	0.6000	air	0.3000
Bear one's hair out	0.9000	bear	0.6000	tear	0.3000
Hightail it (out of bear	0.9000	bear	0.6000	there	0.3000
Not a bear out of place	0.9000	bear	0.6000	hair	0.3000
If the shoe fits, bear it	0.9000	bear	0.6000	wear	0.3000
Bear the living daylights out of someone	0.9000	bear	0.6000	scare	0.3000
Lion's bear	0.9000	bear	0.6000	share	0.3000
Put your money bear your mouth is	0.9000	bear	0.6000	where	0.3000
Bear the rod and spoil the child	0.9000	bear	0.6000	spare	0.3000
Back to bear one	0.9000	bear	0.6000	square	0.3000
Not have a bear	0.9000	bear	0.6000	prayer	0.3000
Give someone a run for bear money	0.9000	bear	0.6000	their	0.3000
Bear one's hair out	0.9000	bear	0.6000	tear	0.3000
Turnabout is bear play	0.9000	bear	0.6000	fair	0.3000
Up, up in the bear	0.9000	bear	0.6000	air	0.3000
Bear the rod and spoil the child	0.9000	bear	0.6000	spare	0.3000
Lion's bear	0.9000	bear	0.6000	share	0.3000
If the shoe fits, bear it	0.9000	bear	0.6000	wear	0.3000
Bear the living daylights out of someone	0.9000	bear	0.6000	scare	0.3000
Not a bear out of place	0.9000	bear	0.6000	hair	0.3000

Experiments in Captioning, *Wit a Twist*

The Amusemater Captioner



Experiments in Captioning, Wit a Twist

The Amusemater Captioner



Experiments in Captioning, *Wit* a Twist

The Amusemater Captioner



Experiments in Captioning, *Wit* a Twist

The Amusemater Captioner



Experiments in Captioning, Wit a Twist

The Amusemater Captioner

select a picture from IMAGENET
285
1 5500

Exploring more more more potential captions 1

Exploring more more more potential captions 1



any tom, desk or harry

Experiments in Captioning, *Wit a Twist*

The Amusemater Captioner



Put your money chair your mouth is

Experiments in Captioning, *Wit a Twist*

The Amusemater Captioner



Harp as a tack

Experiments in Captioning, *Wit* a Twist

The Amusemater Captioner

select a picture from IMAGENET
1
1 5500

Exploring more potential captions

Exploring more and more and more and more potential captions



look what the cat dragonfly in

Experiments in Captioning, *Wit* a Twist

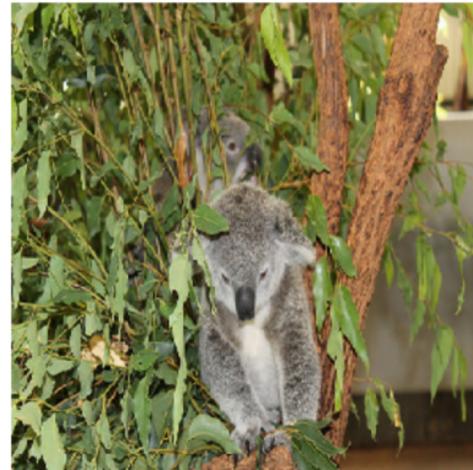
The Amusemater Captioner

select a picture from IMAGENET



Exploring more and more and more potential captions

Exploring more and more and more potential captions



Trying to koala a fly with an elephant gun

Experiments in Captioning, Wit a Twist

The Amusemater Captioner

select a picture from IMAGENET
1 3503 5500

Exploring more potential captions

Exploring more and more and more potential captions



Brambling up is hard to do

Experiments in Captioning, *Wit* a Twist

The Amusemater Captioner

select a picture from IMAGENET



Exploring more and more and more and more potential captions



You will be in seventh heron

Show other captions

Experiments in Captioning, *Wit a Twist*

The Amusemater Captioner

select a picture from IMAGENET

1

5086

5500

Exploring more and more and more and more potential captions



shutterstockphoto

Time to bury the helmet

Experiments in Captioning, Wit a Twist

The Amusemater Captioner

select a picture from IMAGENET

3680

1 550

Exploring more more potential captions

Exploring more more more potential captions



kill the goose that laid the goldfish egg

1500

idioms

well it is a bit much
another bite at the cherry
busy as a bee
they say a cat has nine lives
a cat in gloves catches no mice
time for a cat nap
it was a cold day in july
a bit of a cold fish
a cut above
a cut below
always a day late and a dollar short
being a dog in the manger
we are a few sandwiches short of a picnic
never missing out on a good deal
it means a great deal
offering to be a guinea pig
within a hair's breadth
hombody, homester, a home bird
giving a hundred and ten percent
now just a lame duck
give a leg up
a bit of a lemon
taking on a life of its own
a little bird told me

Senti Word Net corpus

taking place by degrees

happening without warning or in a short space of time; "a sudden storm"; "a sudden decision"; "a sudden cure"
exceedingly sudden and unexpected; "came to an abrupt stop"; "an abrupt change in the weather"
marked by abrupt transitions; "choppy prose"

occurring unexpectedly and requiring urgent action; "emergent repair of an aorta"
sudden and loud; "an explosive laugh"

sudden and severe; "fulminant pain"; "fulminant fever"

very sudden and in great amount or degree; "a sharp drop in the stock market"
(of a topographical gradient) not steep or abrupt; "a gradual slope"
marked by moderate steepness; "an easy climb"; "a gentle slope"

having a slanting form or direction; "an area of gently sloping hills"; "a room with a sloping ceiling"

having a sharp inclination; "the steep attic stairs"; "steep cliffs"

extremely steep; "an abrupt canyon"; "the precipitous rapids of the upper river"; "the precipitous hills of Chinese p
very steep; having a prominent and almost vertical front; "a bluff headland"; "where the bold chalk cliffs of Englan
sharply inclined; "a heavy grade"

extremely steep; "the great perpendicular face of the cliff"

somewhat steep

having very steep sides

conforming to the rules of grammar or usage accepted by native speakers; "spoke in grammatical sentences"

not grammatical; not conforming to the rules of grammar or accepted usage

(of a word or expression) not agreeing with grammatical principles

feeling or showing gratitude; "a grateful heart"; "grateful for the tree's shade"; "a thankful smile"

feeling or expressive of gratitude; "was appreciative of his efforts"; "an appreciative word"

feeling happy appreciation; "glad of the fire's warmth"

not feeling or showing gratitude; "ungrateful heirs"; "How sharper than a serpent's tooth it is / To have a thankless

not feeling or expressing gratitude; "unappreciative of nature's bounty"

of a cell or organism having a single set of chromosomes

of a cell or organism having two sets of chromosomes or twice the haploid number; "diploid somatic cells"

of a cell or organism having more than twice the haploid number of chromosomes; "a polyploid cell"; "a polyploid

of a cell or organism having three complete sets of chromosomes; "human triploid fetuses are usually spontaneou

enjoying or showing or marked by joy or pleasure; "a happy smile"; "spent many happy days on the beach"; "a hap

Preprocessing and precomputation



Scraped and preprocessed idioms



Precomputed lookup dictionaries for fast access to rhymes and to phrases



Precomputed a 55K x 55K correlation matrix of TF-IDF for contextually ‘nearby’ words in SentiWordNet



Designed, developed, and tested, Streamlit app

Interactive app algorithm



Classify (label) the image



Generate a semantic family by finding words semantically related to the image label



Generate a phonetic family (rhymes, assonances, etc) for each semantically related word.
Compute and rank their proximity to their parent



For each word in the phonetic families, find phrases in the idiom list containing the word



Rank the phrases (depending on user settings)



Substitute the phonetically related word with its parent and caption the image



Display in Streamlit app

APIs and pretrained models used

Beautiful Soup

Xception pre-trained neural for image classification

NLTK, Gensim, Sklearn, and misc. phonetics packages

Streamlit

Contextually 'nearby' words (*not* synonyms)

Not yet
implemented

cliff ['steep', 'taluses', 'jutted']
panda ['ailuropodidae', 'giant', 'classifications']
lampshade ['supports', 'curved', 'pair']
goose ['brant', 'honking', 'greylag']
submarine ['exiting', 'submarined', 'undersea']
crate ['glassed', 'uncrate', 'packing']
fox ['reynard', 'vulpine', 'hounds']
bag ['drawstring', 'capacious', 'burlap']
dowitcher ['breast', 'grey', 'red']
flute ['fipple', 'incarnated', 'tabor']
chickadee ['blackcap', 'crown', 'dark']
sock ['argyle', 'reaches', 'sandals']
microwave ['resonances', 'maser', 'diathermy']
loaf ['flaked', 'bread', 'figurine']
bobsled ['bobbed', 'riding', 'screaming']
capuchin ['douroucouli', 'saki', 'uakari']
zebra ['mussels', 'resembled', 'leopards']
dish ['served', 'cantones', 'rice']
streetcar ['trundled', 'avenue', 'operator']
tissue ['connective', 'necrotic', 'mesodermal']
dog ['breed', 'greyhound', 'trained']