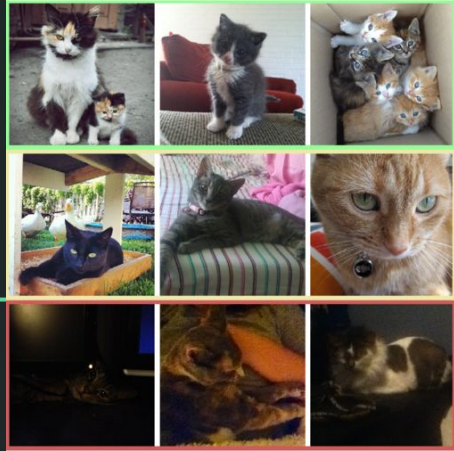
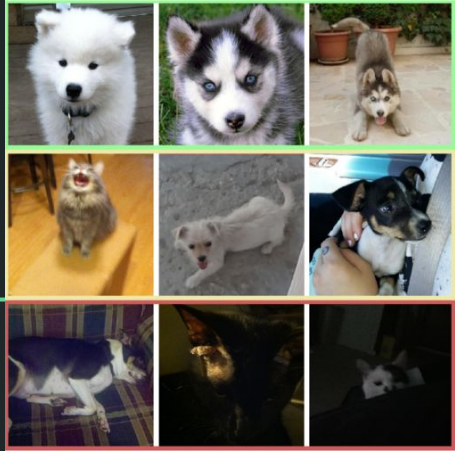
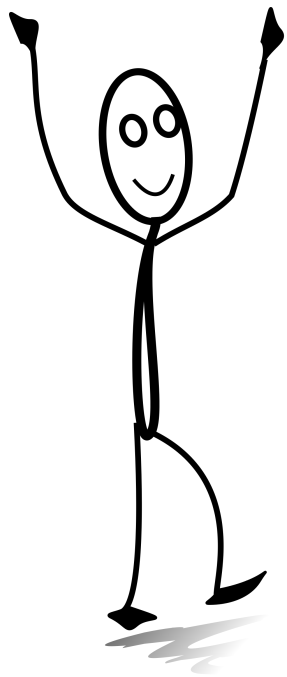
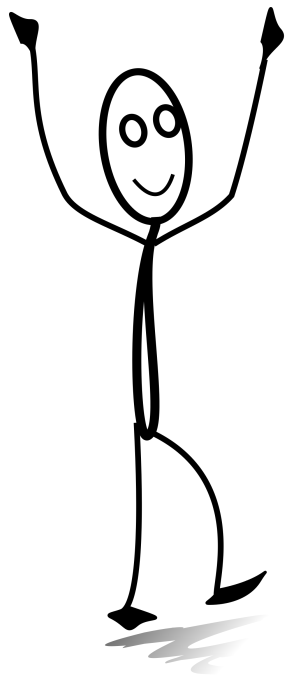


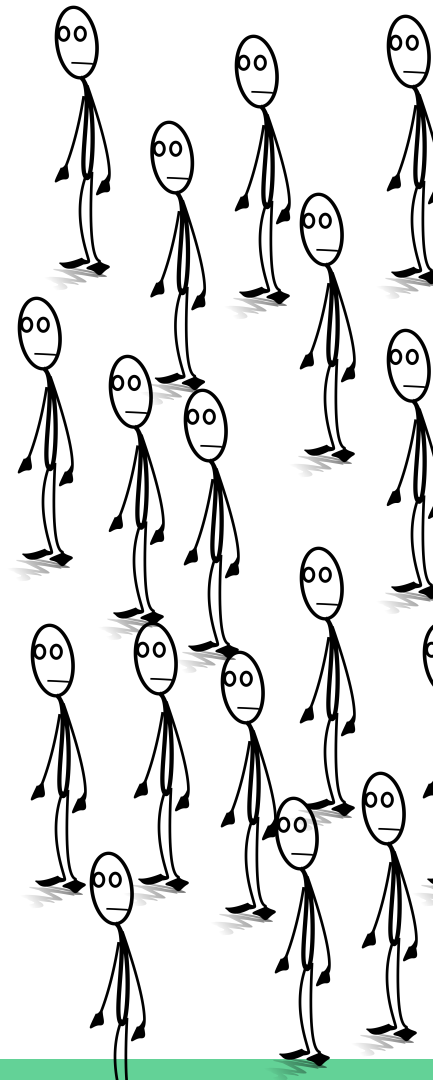
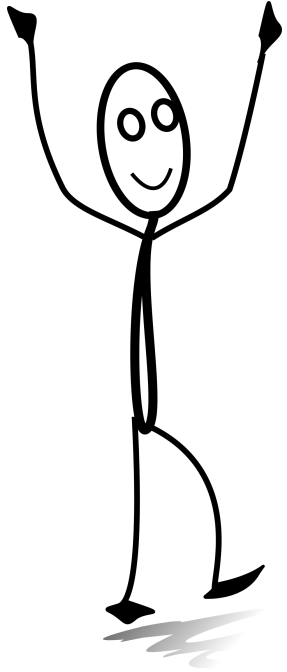
# Cats and Captions vs. Creators and the Clock: *Comparing Multimodal Content to Context in Predicting Relative Popularity*

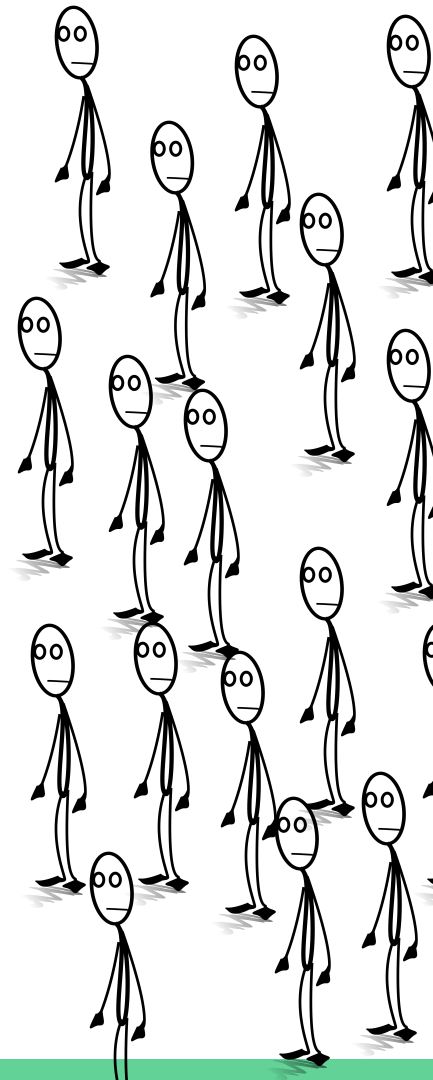
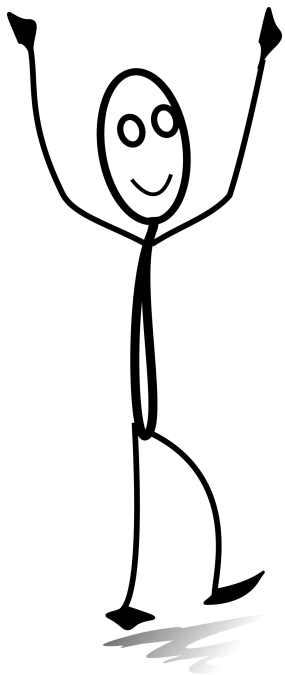


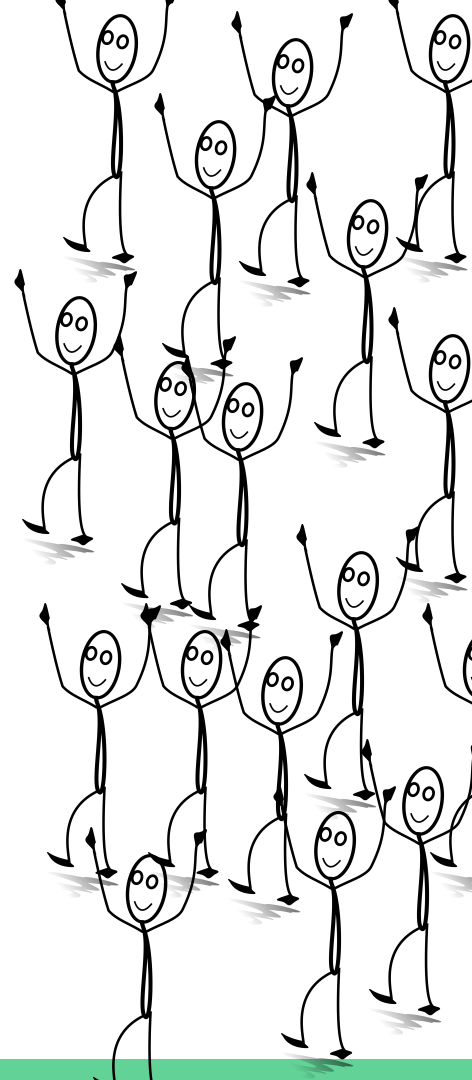
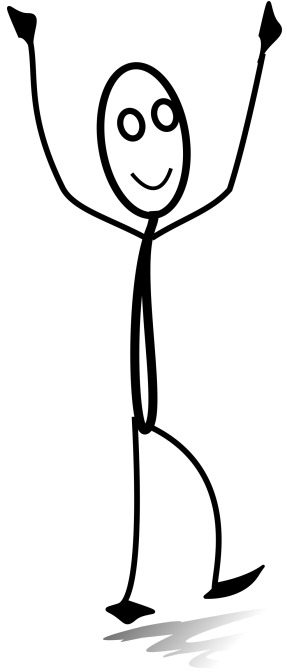
**Jack Hessel**, Lillian Lee, David Mimno  
Cornell University

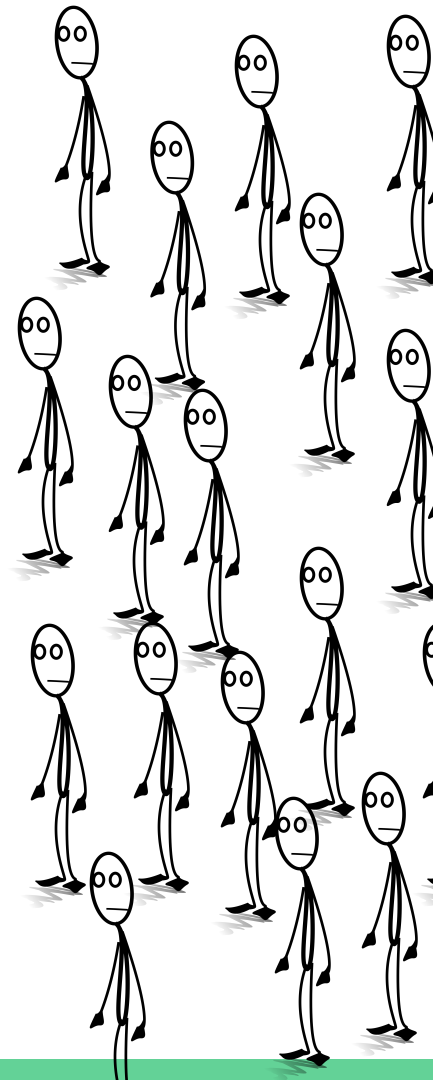
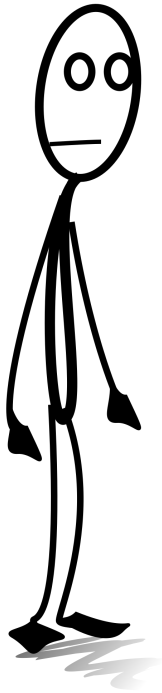


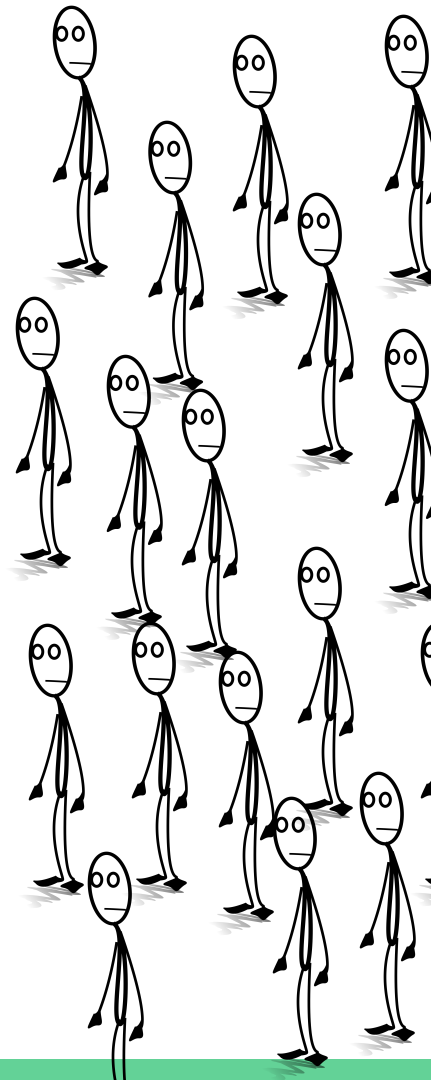
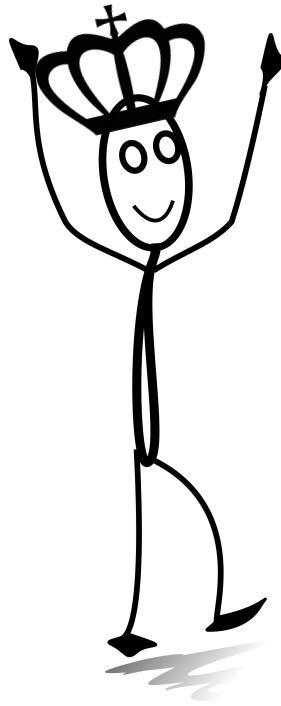




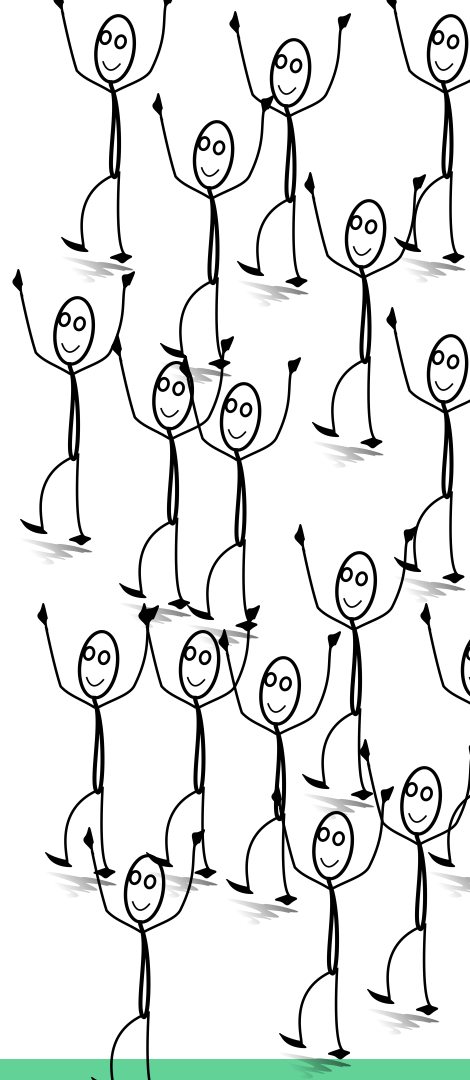
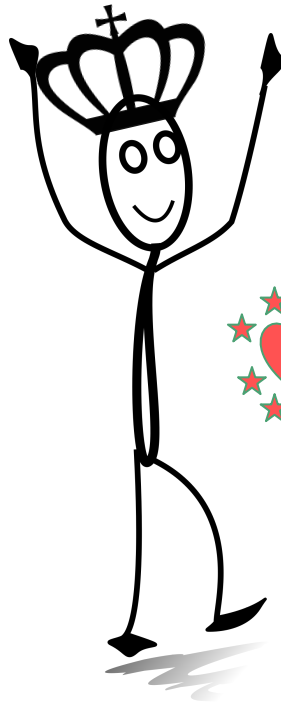


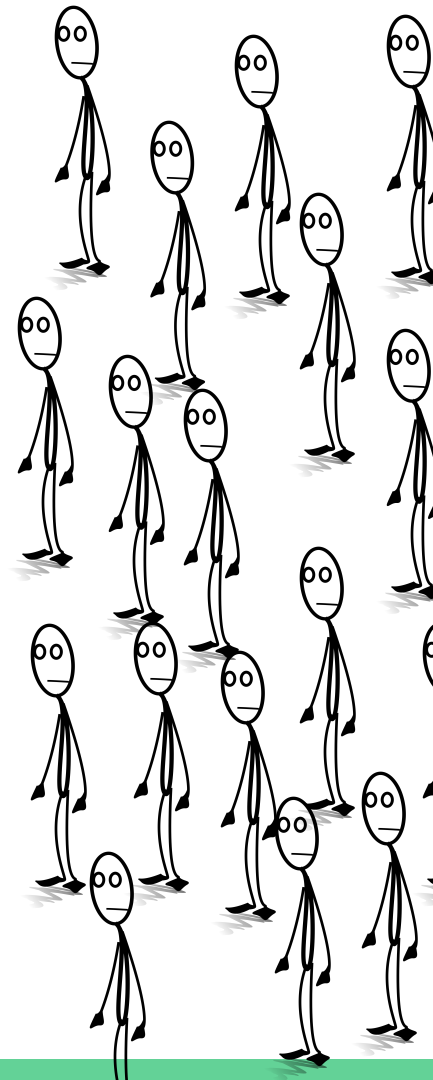








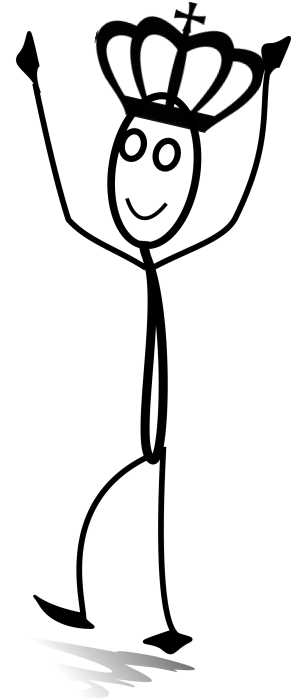






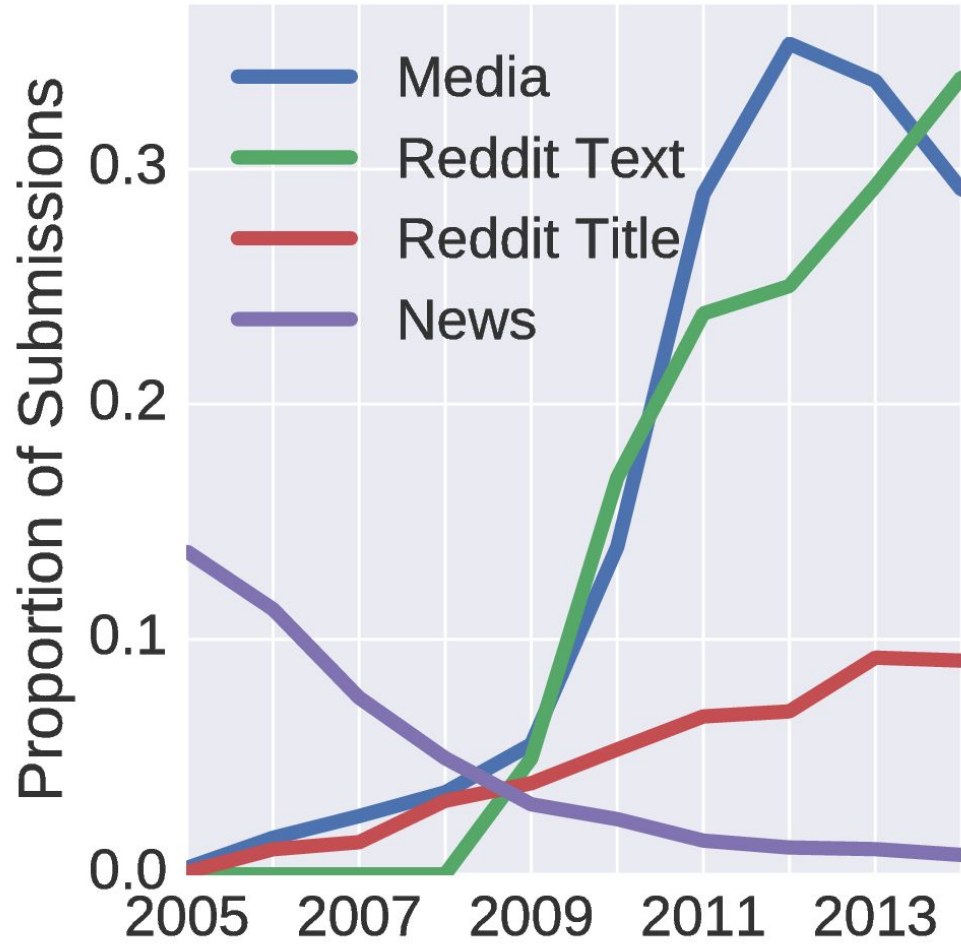


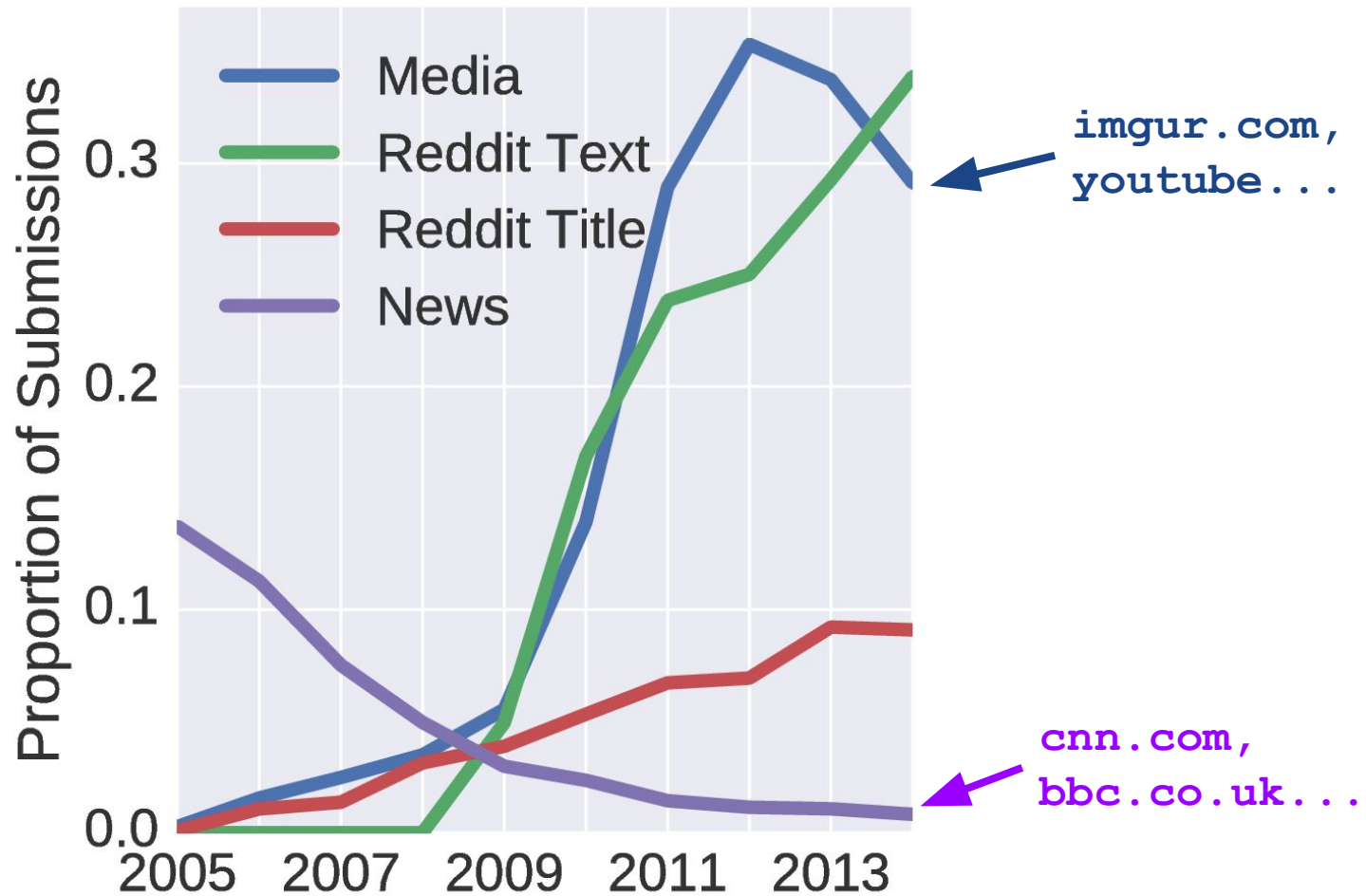
VS



Can we isolate the  
effects of  
*content*  
rather than context?







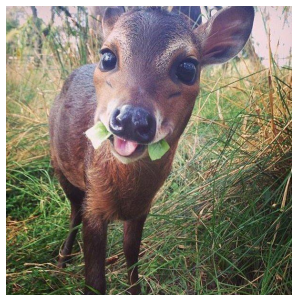




"Tonight, I carved a pumpkin. I also doused it in lighter fluid and lit it on fire." - /r/pics



"You have to go to the border for food Fish Tacos [San Diego]" - /r/FoodPorn



"Snacks!" - /r/aww

	# Users	#/% Img	Cap Len
pics	2108K	2472K/70%	9.84
aww	1010K	954K/81%	9.13
cats	109K	100K/73%	8.97
MakeupAddiction (MA)	77K	58K/57%	13.67
FoodPorn (FP)	74K	50K/77%	9.39
RedditLaqueristas (RL)	27K	39K/73%	11.12



"Glamor Leaves" - /r/RedditLaqueristas



The grass is always  
greener



This is why you get  
two cats



The grass is always  
greener



This is why you get  
two cats

# Complication 1: Rich get richer



# Complication 2: Identity Matters



LINKS FROM: PAST MONTH ▾

YOU ARE NOT A MEMBER OF THIS COMMUNITY. PLEASE RESPECT THAT BY NOT DOWNVOTING

4040



A little late but I did a Frida look for Halloween I'm pretty proud of! [i.redd.it](#)

submitted 27 days ago by [osaosa](#) [IG: selenaaasx](#)

[83 comments](#) [share](#) [save](#) [hide](#) [report](#)

3878

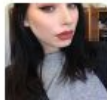


Bat makeup for halloween :) [i.reddituploads.com](#)

submitted 29 days ago by [sephorv](#)

[44 comments](#) [share](#) [save](#) [hide](#) [report](#)

3813

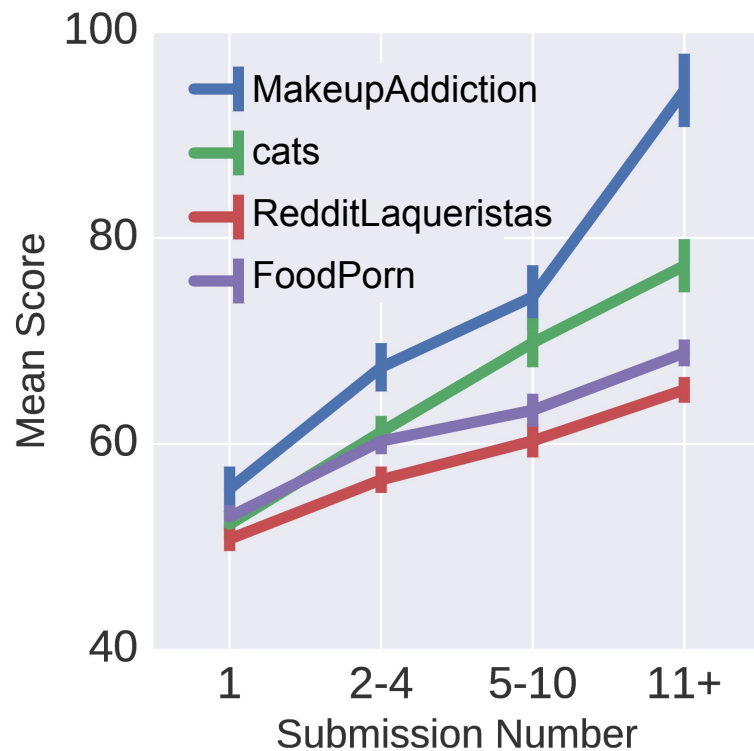


I can conquer the world in this makeup.

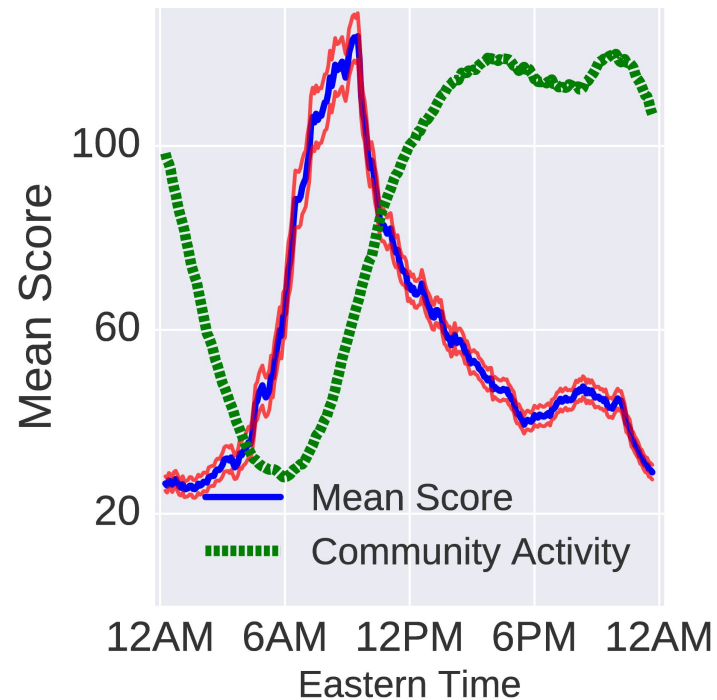
[i.reddituploads.com](#)

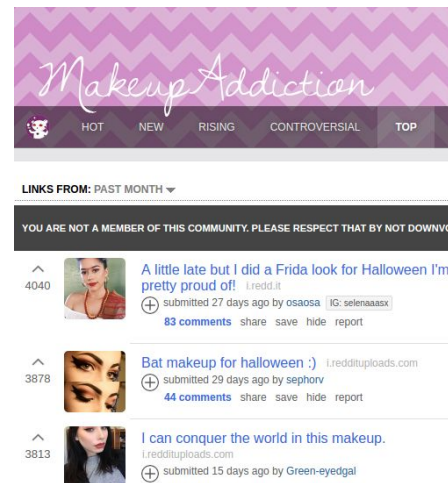
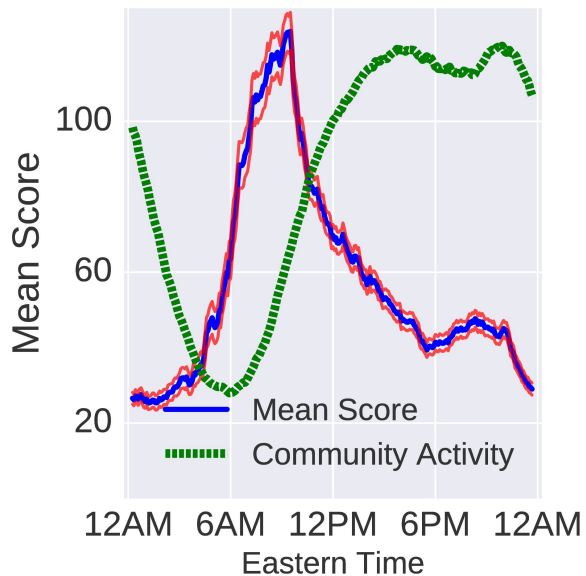
submitted 15 days ago by [Green-eyedgal](#)

## Complication 2: Identity Matters



## Complication 3: Minutes matter









The grass is always  
greener



This is why you get  
two cats



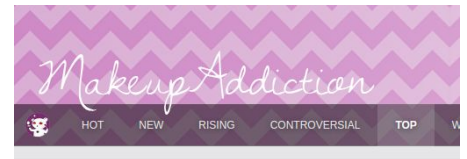
The grass is always greener



This is why you get two cats




*13 Seconds Apart!*

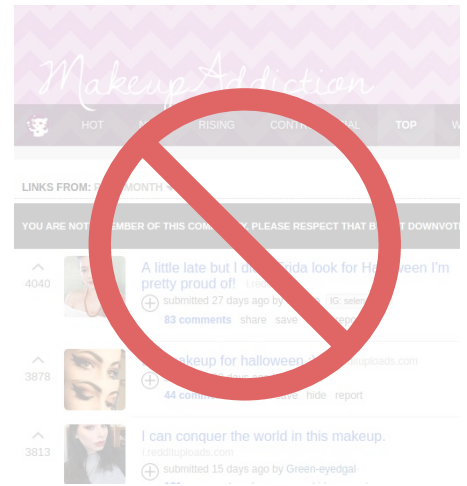
		aww	pics	cats	MA	FP	RL
Timing	Random	50.0	50.0	50.0	50.0	50.0	50.0
	Time	50.2	50.2	50.7	50.4	49.7	50.6

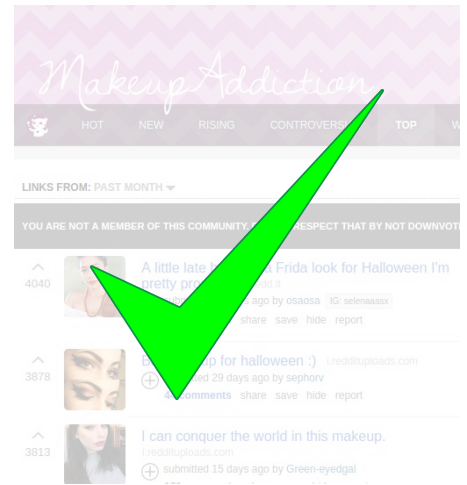


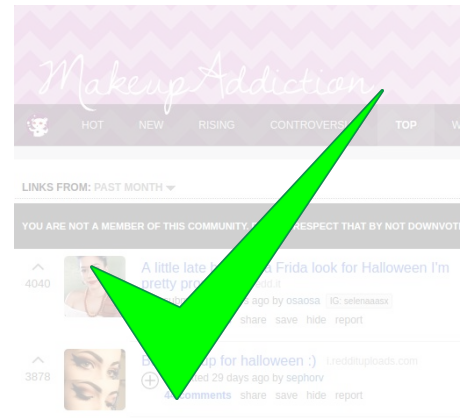
LINKS FROM: PAST MONTH

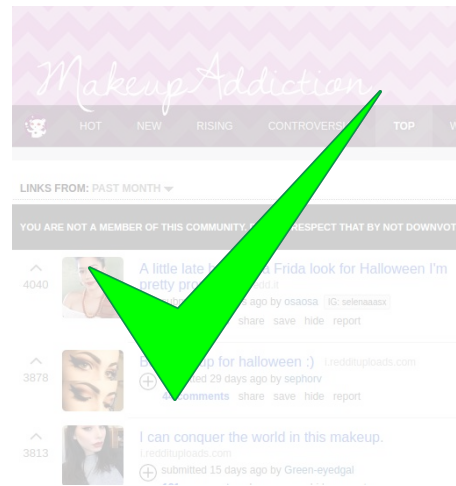
YOU ARE NOT A MEMBER OF THIS COMMUNITY. PLEASE RESPECT THAT BY NOT DOWNVOTING

- 4040  **A little late but I did a Frida look for Halloween I'm pretty proud of!** 1 redd.it  
 submitted 27 days ago by [osaosa](#) IG: selenaaax  
**83 comments** share save hide report
- 3878  **Bat makeup for halloween :)** 1 reddituploads.com  
 submitted 29 days ago by [sephorv](#)  
**44 comments** share save hide report
- 3813  **I can conquer the world in this makeup.** 1 reddituploads.com  
 submitted 15 days ago by [Green-eyedgal](#)



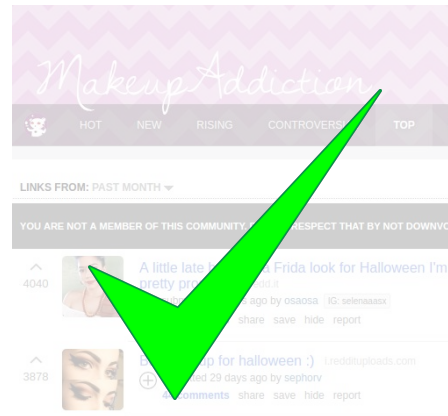








		aww	pics	cats	MA	FP	RL
Timing	Random	50.0	50.0	50.0	50.0	50.0	50.0
	Earlier	<i>51.7</i>	<i>51.1</i>	<i>49.9</i>	48.9	<i>48.6</i>	48.7
	Time	50.2	50.2	<i>50.7</i>	<i>50.4</i>	<i>49.7</i>	<i>50.6</i>





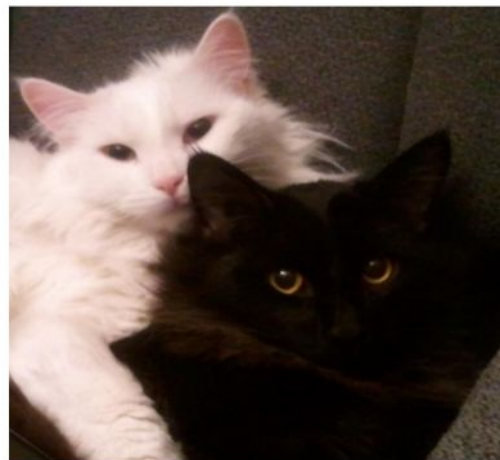
The grass is always  
greener



This is why you get  
two cats



The grass is always greener



This is why you get two cats

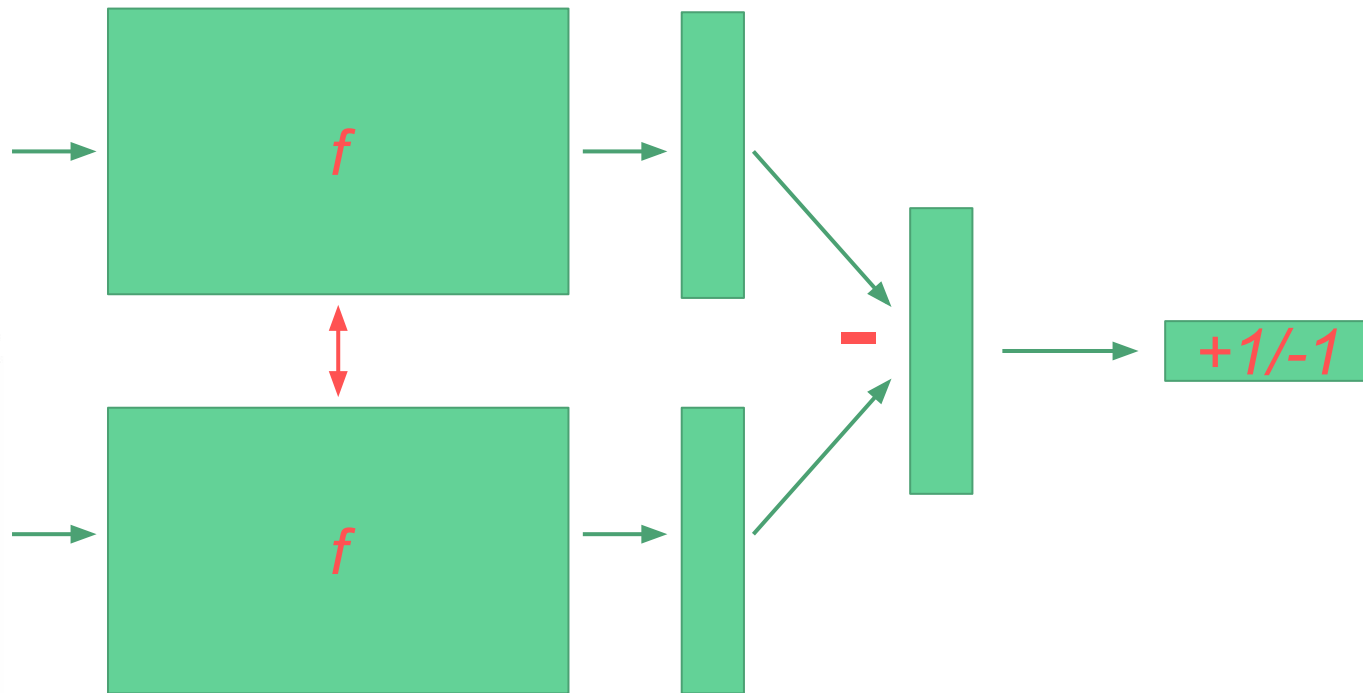
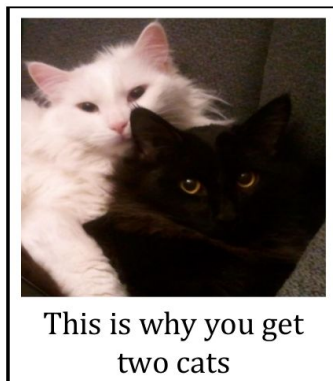
---

	aww	pics	cats	MA	FP	RL
Humans	60.0	63.6	59.6	62.2	72.7	67.2

---

	Max/Avg Win	Med/Avg Diff	# Pairs
pics	30/15 sec	117/478	44K
aww	30/15 sec	90/393	33K
cats	15/7 min	69/231	15K
MA	60/24 min	88/227	10K
FP	120/53 min	62/188	8K
RL	30/14 min	56/118	9K

# Machine learning experiments



# Unimodal Results (cross-val)

# Unimodal Results (cross-val)

	aww	pics	cats	MA	FP	RL
Type	50.6	51.2	50.7	52.8	51.8	56.1
Activity	51.1	53.6	52.8	55.0	53.9	60.6
Quality	54.7	55.5	52.9	60.7	55.5	<b>67.3</b>
Struct	56.2	54.8	56.5	50.9	52.3	52.5
Topic	55.2	55.8	56.8	60.4	55.2	55.5
DAN	58.6	58.3	58.5	62.2	57.6	59.8
LSTM	59.4	58.8	58.7	61.0	57.0	59.1
Bi-LSTM	59.7	58.9	59.3	61.8	57.8	59.6
Unigram	59.7	58.6	59.5	63.0	57.6	60.8
HOG	51.7	52.8	51.9	53.5	53.5	53.5
GIST	52.7	53.0	53.5	55.9	56.5	56.3
ColorHist	55.3	53.7	55.6	55.0	56.5	54.5
VGG-19	63.4	58.9	61.1	62.4	62.8	62.1
ResNet50	<b>64.8</b>	<b>60.0</b>	<b>62.6</b>	<b>64.9</b>	<b>65.2</b>	64.2

**User Features**

**Text Features**

**Image Features**



# Unimodal Results (cross-val)

	aww	pics	cats	MA	FP	RL
Type	50.6	51.2	50.7	52.8	51.8	56.1
Activity	51.1	53.6	52.8	55.0	53.9	60.6
Quality	54.7	55.5	52.9	60.7	55.5	<b>67.3</b>
Struct	56.2	54.8	56.5	50.9	52.3	52.5
Topic	55.2	55.8	56.8	60.4	55.2	55.5
DAN	58.6	58.3	58.5	62.2	57.6	59.8
LSTM	59.4	58.8	58.7	61.0	57.0	59.1
Bi-LSTM	59.7	58.9	59.3	61.8	57.8	59.6
Unigram	59.7	58.6	59.5	63.0	57.6	60.8
HOG	51.7	52.8	51.9	53.5	53.5	53.5
GIST	52.7	53.0	53.5	55.9	56.5	56.3
ColorHist	55.3	53.7	55.6	55.0	56.5	54.5
VGG-19	63.4	58.9	61.1	62.4	62.8	62.1
ResNet50	<b>64.8</b>	<b>60.0</b>	<b>62.6</b>	<b>64.9</b>	<b>65.2</b>	64.2

**User Features**

**Text Features**

**Image Features**

# Multimodal Results

# Multimodal Results (crossval)

## Multimodal Results (crossval)

	aww	pics	cats	MA	FP	RL
Time + User	54.1	54.7	52.1	58.8	54.2	64.8
All User	56.3	55.3	54.6	60.9	56.0	<b><u>68.4</u></b>
ResNet50	64.8	60.0	62.6	64.9	65.2	64.2
Text + Image	<b><u>67.1</u></b>	<b><u>62.7</u></b>	<b><u>65.9</u></b>	<b><u>67.7</u></b>	<b><u>65.8</u></b>	66.4

# Multimodal Results (crossval)

	aww	pics	cats	MA	FP	RL
Time + User	54.1	54.7	52.1	58.8	54.2	64.8
All User	56.3	55.3	54.6	60.9	56.0	<b><u>68.4</u></b>
ResNet50	64.8	60.0	62.6	64.9	65.2	64.2
Text + Image	<b><u>67.1</u></b>	<b><u>62.7</u></b>	<b><u>65.9</u></b>	<b><u>67.7</u></b>	<b><u>65.8</u></b>	66.4

Best unimodal



# Multimodal Results (crossval)

	aww	pics	cats	MA	FP	RL
Time + User	54.1	54.7	52.1	58.8	54.2	64.8
All User	56.3	55.3	54.6	60.9	56.0	<b><u>68.4</u></b>
ResNet50	64.8	60.0	62.6	64.9	65.2	64.2
Text + Image	<b><u>67.1</u></b>	<b><u>62.7</u></b>	<b><u>65.9</u></b>	<b><u>67.7</u></b>	<b><u>65.8</u></b>	66.4

Best unimodal



*Multimodal  
beats unimodal!*



# Multimodal Results (crossval + fully heldout)

	aww	pics	cats	MA	FP	RL
Time + User	54.1	54.7	52.1	58.8	54.2	64.8
All User	56.3	55.3	54.6	60.9	56.0	<b><u>68.4</u></b>
ResNet50	64.8	60.0	62.6	64.9	65.2	64.2
Text + Image	<b><u>67.1</u></b>	<b><u>62.7</u></b>	<b><u>65.9</u></b>	<b><u>67.7</u></b>	<b><u>65.8</u></b>	66.4

Best unimodal

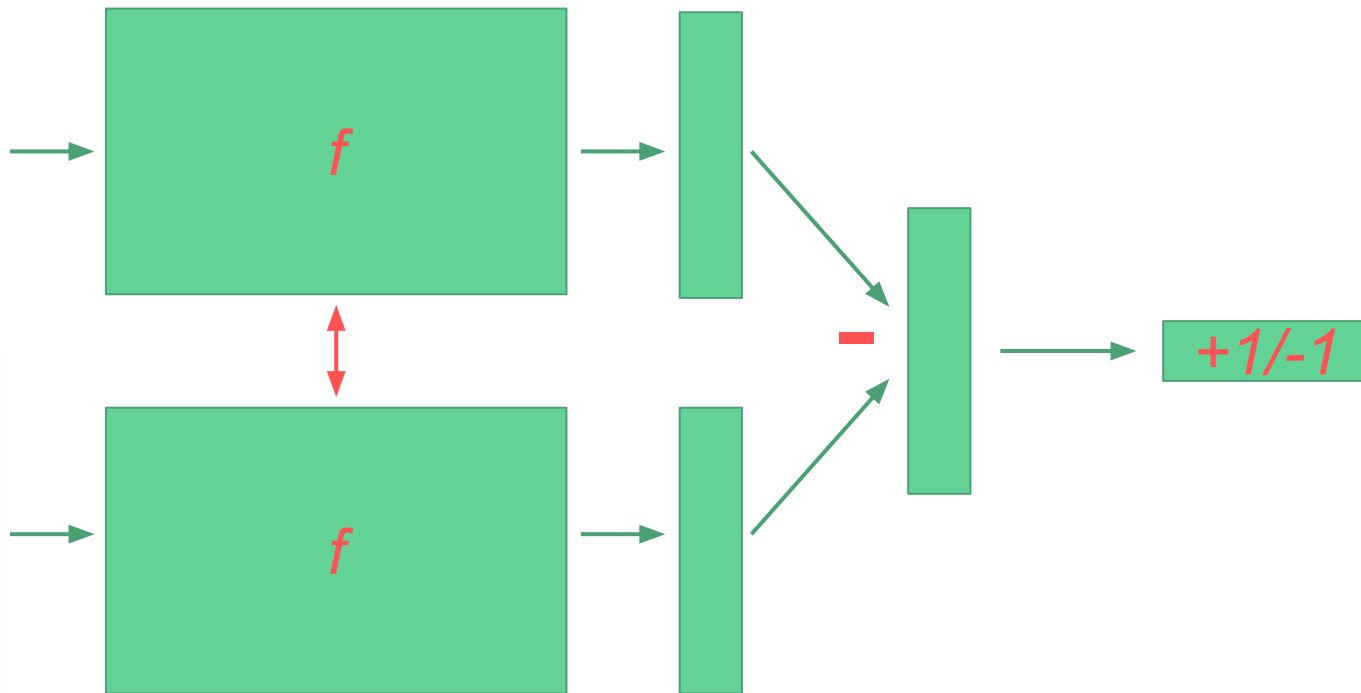
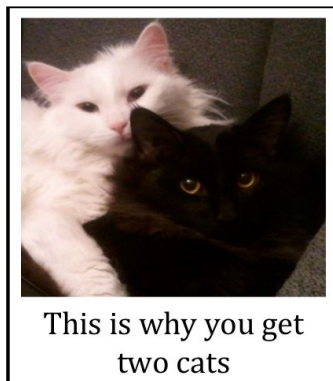


Multimodal  
beats unimodal!



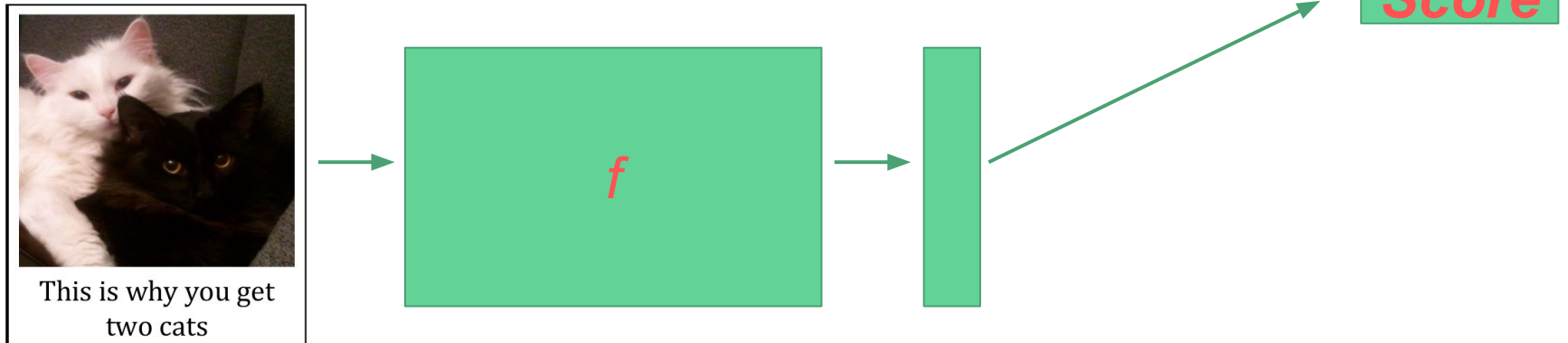
	aww	pics	cats	MA	FP	RL
Time + User	55.5	51.7	52.6	56.9	52.8	60.5
All User	60.4	51.0	54.3	<b>63.1</b>	57.9	<b>66.0</b>
Text + Image	<b>65.5</b>	<b>66.0</b>	<b>67.3</b>	62.7	<b>62.6</b>	65.4

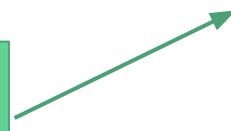
# Machine learning experiments



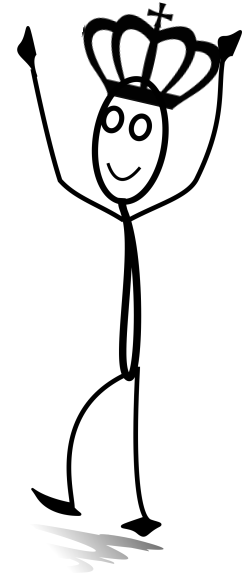
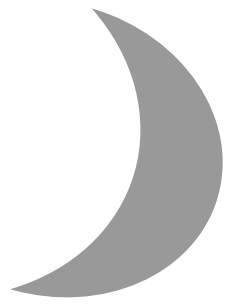


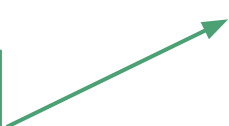
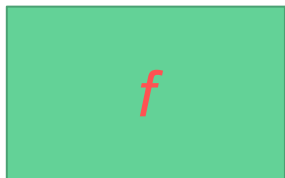
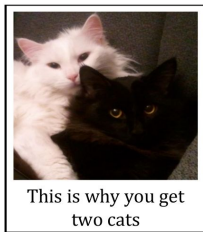
# Machine learning experiments





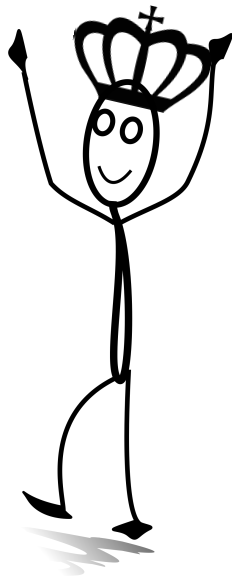
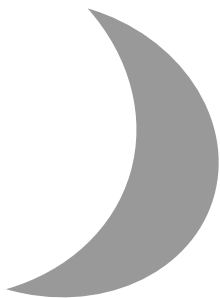
Score



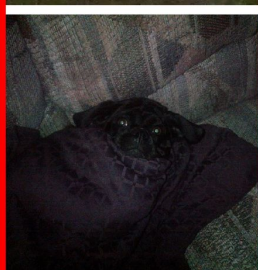
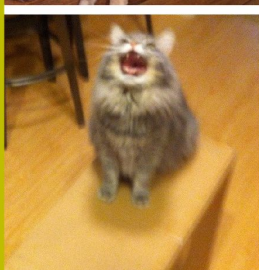
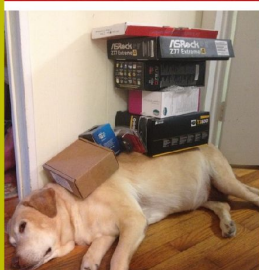


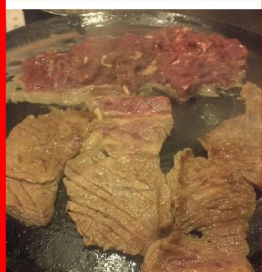
Score

Promote quality  
content posted at  
the wrong time!



Promote quality  
content posted by  
non-celebrities!



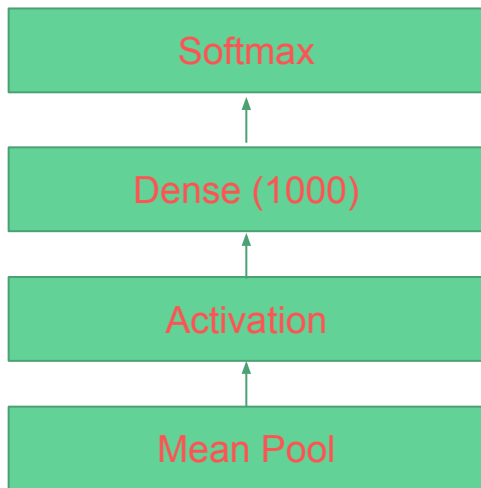




IM  GENET



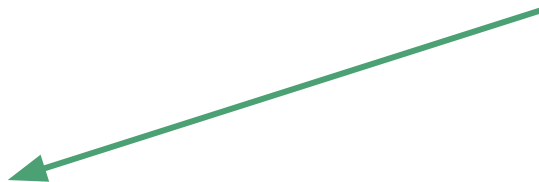
IMAGENET





IMAGENET

**Lots of types of animals!**



Softmax

Dense (1000)

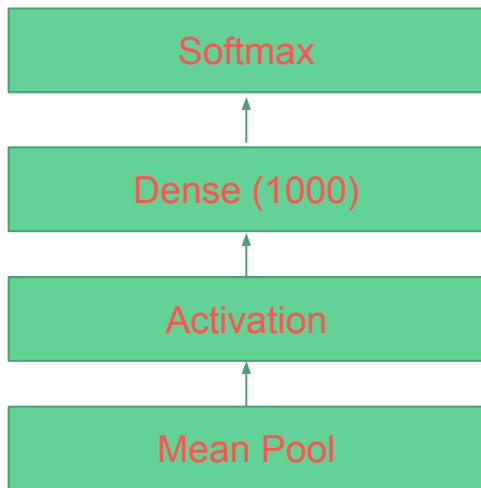
Activation

Mean Pool



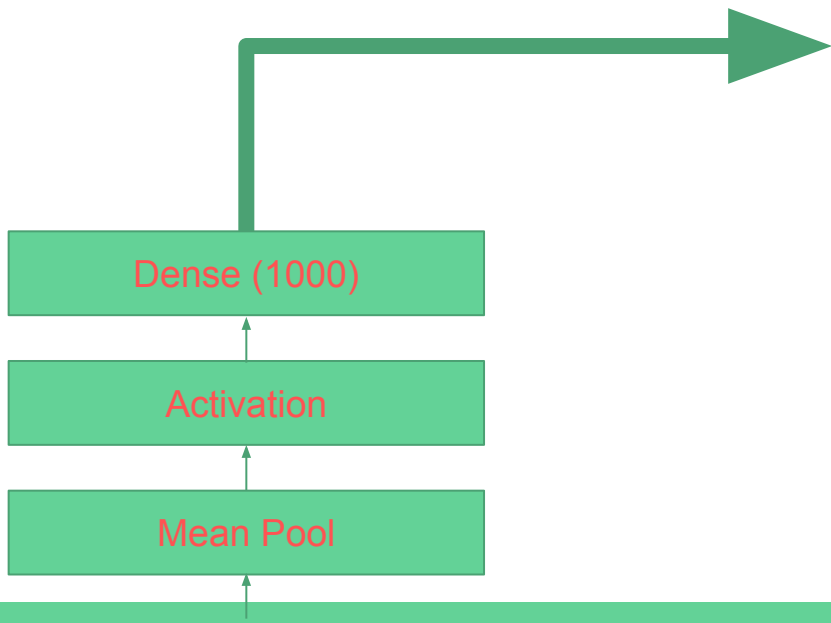


IMAGENET



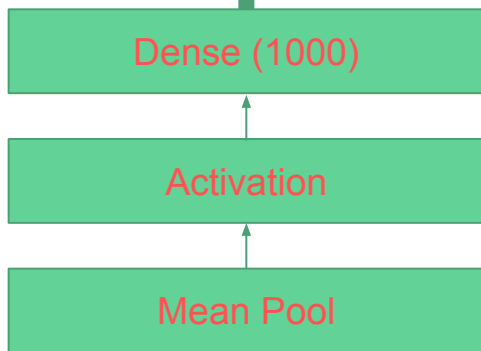


IMAGENET





IMAGENET



```
golden_retriever +0.2290 ***
dingo            +0.2126 ***
Labrador_retriever +0.1960 ***
worm_fence      +0.1864 ***
cheetah         +0.1851 ***
Tibetan_mastiff +0.1830 ***
...
Scotch_terrier  -0.2193 ***
bassinet       -0.2196 ***
wardrobe       -0.2231 ***
miniature_schnauzer -0.2343 ***
four-poster    -0.2841 ***
mosquito_net   -0.2936 ***
```

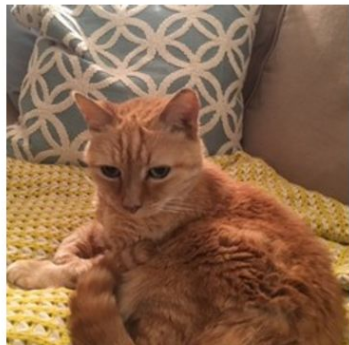
(Significant after applying bonferroni correction)

# Pretrained Models and Data Available!



[www.cs.cornell.edu/~jhessel/cats/cats.html](http://www.cs.cornell.edu/~jhessel/cats/cats.html)

# Pretrained Models and Data Available!



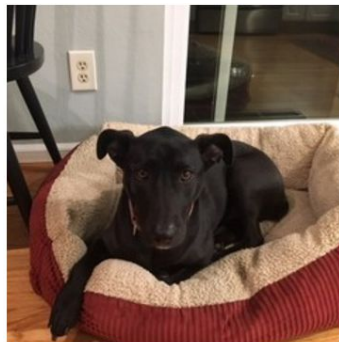
*My 20 year old little girl. Her sister p*

- Score = 87.5/100



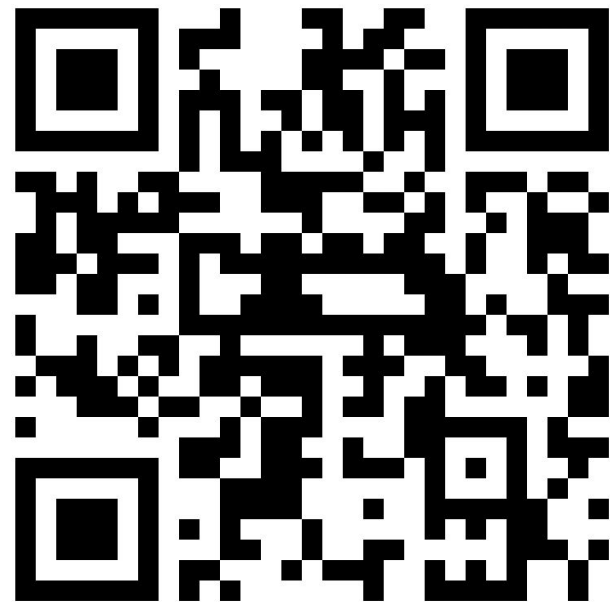
*Please don't sit on me!*

- Score = 27.4/100



*Who says bulldogs can't be cute???*

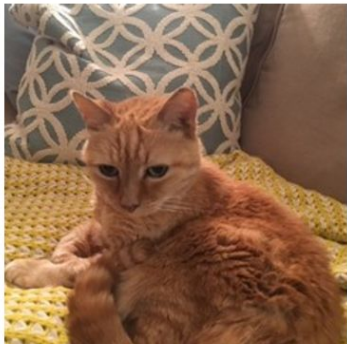
- Score = 13.8/100



[www.cs.cornell.edu/~jhessel/cats/cats.html](http://www.cs.cornell.edu/~jhessel/cats/cats.html)

# Pretrained Models and Data Available!

*lizzy*



*My 20 year old little girl. Her sister p*

- Score = 87.5/100

*taz*



*Please don't sit on me!*

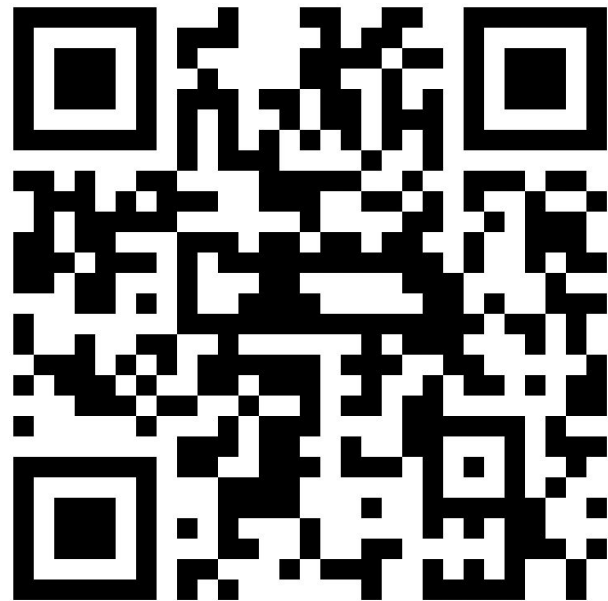
- Score = 27.4/100

*bodhi*



*Who says bulldogs can't be cute???*

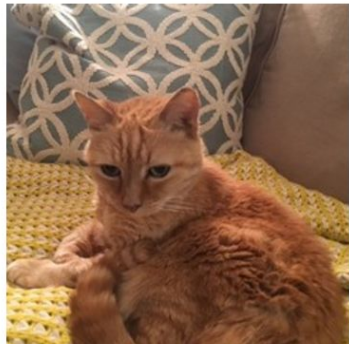
- Score = 13.8/100



[www.cs.cornell.edu/~jhessel/cats/cats.html](http://www.cs.cornell.edu/~jhessel/cats/cats.html)

# Pretrained Models and Data Available!

*lizzy*



*My 20 year old little girl. Her sister p*

- Score = 87.5/100

*taz*



*Please don't sit on me!*

- Score = 27.4/100

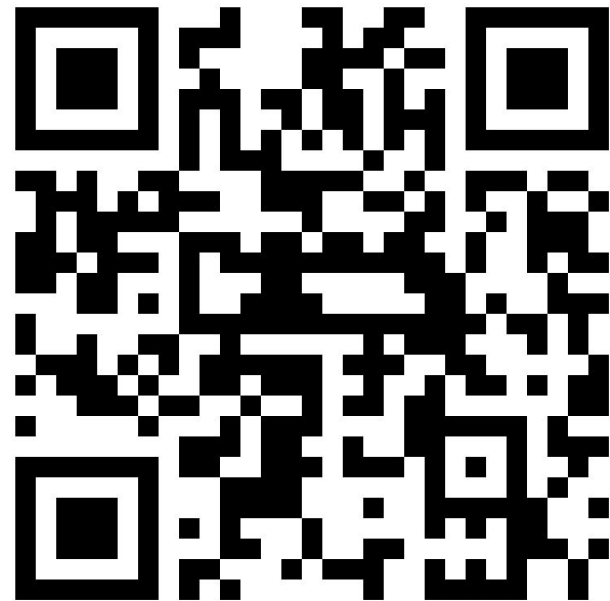
*bodhi*



*Who says bulldogs can't be cute???*

- Score = 13.8/100

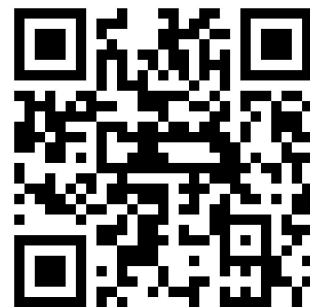
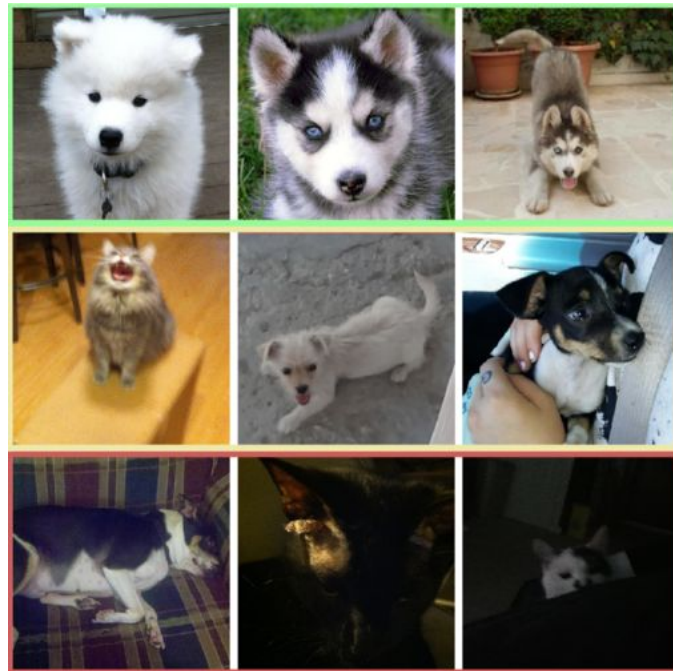
	Max/Avg Win	Med/Avg Diff	# Pairs
pics	30/15 sec	117/478	44K
aww	30/15 sec	90/393	33K
cats	15/7 min	69/231	15K
MA	60/24 min	88/227	10K
FP	120/53 min	62/188	8K
RL	30/14 min	56/118	9K



[www.cs.cornell.edu/~jhessel/cats/cats.html](http://www.cs.cornell.edu/~jhessel/cats/cats.html)

# Thanks to:

- Nvidia for the GPU donations
- Yahoo, Inc.
- National Science Foundation
- Vlad Niculae
- Volunteer human annotators
- Cornell NLP Seminar





# Questions?

<http://www.cs.cornell.edu/~jhessel/>

[jhessel@cs.cornell.edu](mailto:jhessel@cs.cornell.edu)

@jmhessel on Twitter

