Computational Humor

Can a machine have a sense of humor?

Thomas Winters

PhD Student at KU Leuven & FWO Fellow

@thomas_wint

thomaswinters.be

Can a machine have a sense of humor?

Part 1: Humor

Intrinsically human!

Every known human civilisation has some form of humor



Caron, J.E.: From ethology to aesthetics: Evolution as a theoretical paradigm for research on laughter, humor, and other comic phenomena. Humor15(3), 245–281(2002)

Purpose of Humor = Sign of Intelligence?



Brain is rewarding **noticing incongruities** and successfully **resolving them**

\+ **linguistic skills**
(required to create jokes)

= **Evolutionary advantage!**

How does humor work?

Incongruity-Resolution Theory

Setup *Two fish are in a tank. Says one to the other:*

Punchline *"Do you know how to drive this thing?"*

Obvious Interpretation



Hidden Interpretation



Based on: Ritchie, G. (1999). Developing the incongruity-resolution theory.

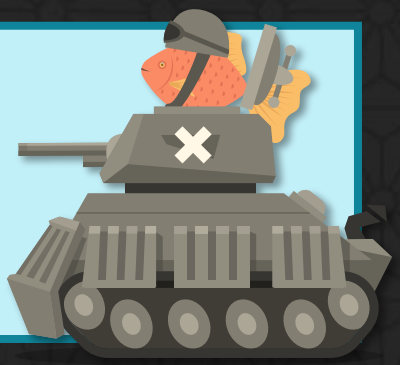
Human-focused definition!

Machine should **not only spot two mental images**

Obvious Interpretation



Hidden Interpretation



But also this is **not too hard or too easy** for a **human**!

Humor = AI-complete?

~ as hard as making computers as “intelligent” as people

Part 2: Text & Humor Generation

Hard-coding a sense of humor

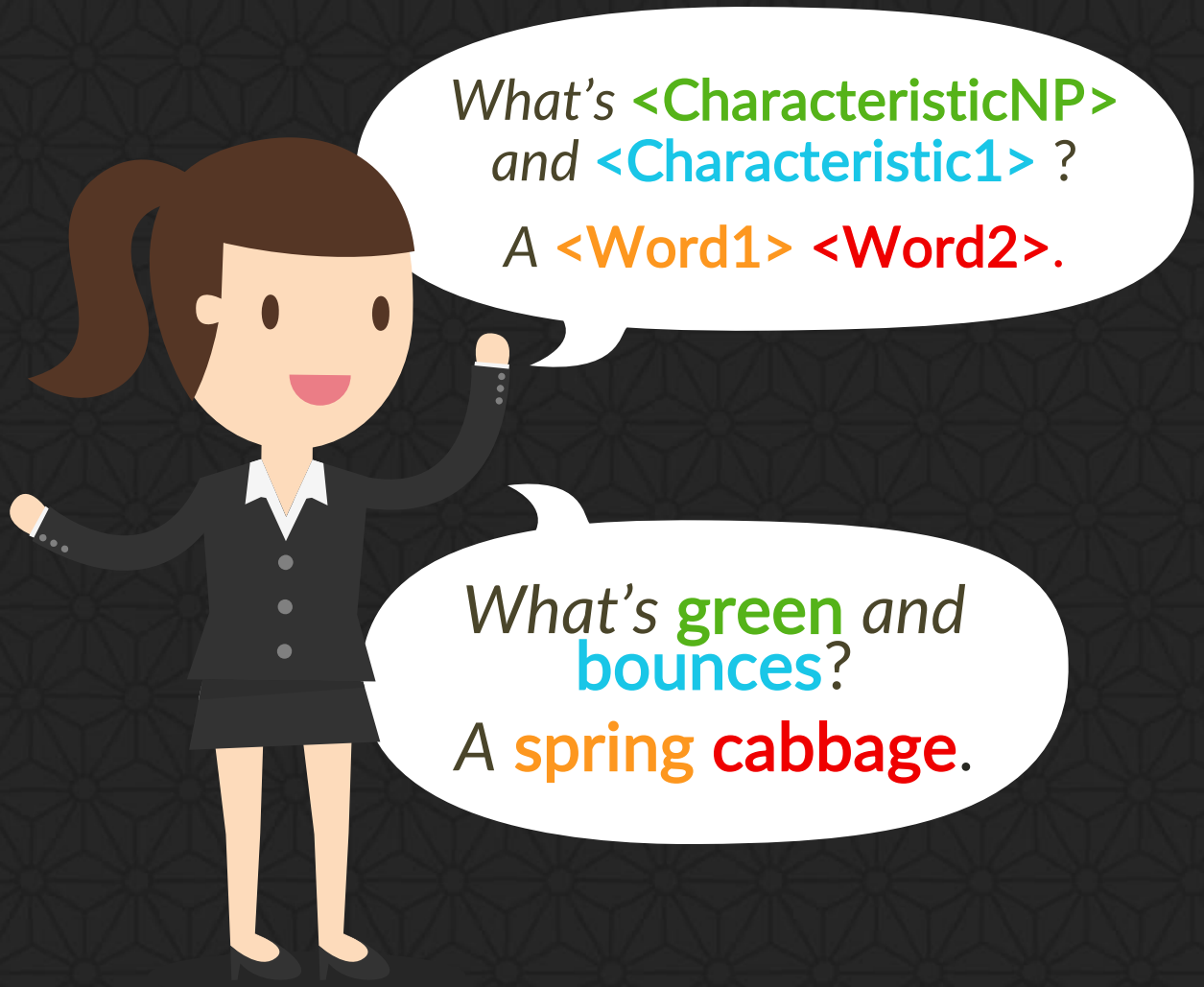
JAPE Joke Examples

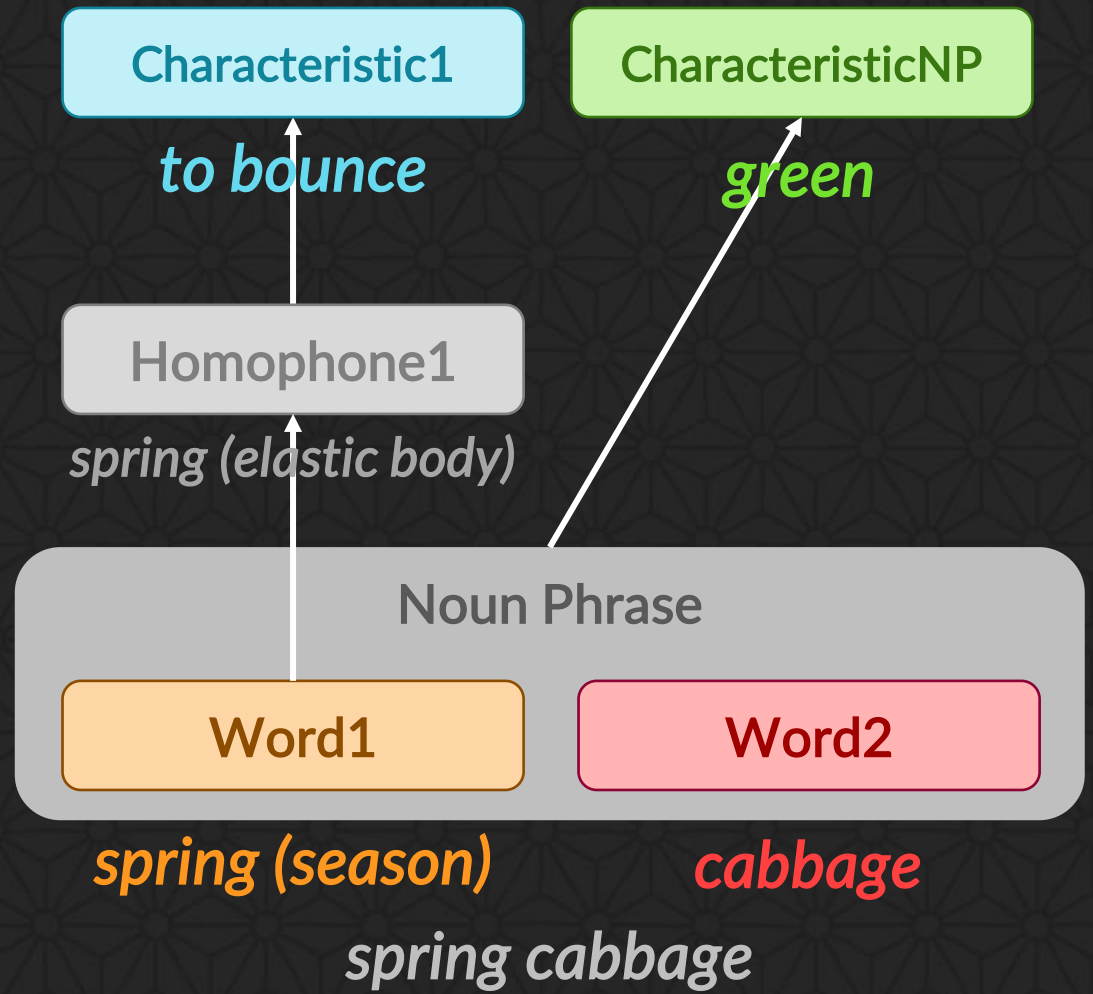
What is the difference between leaves and a car?
One you brush and rake, the other you rush and brake

What do you call a strange market?
A bizarre bazaar.



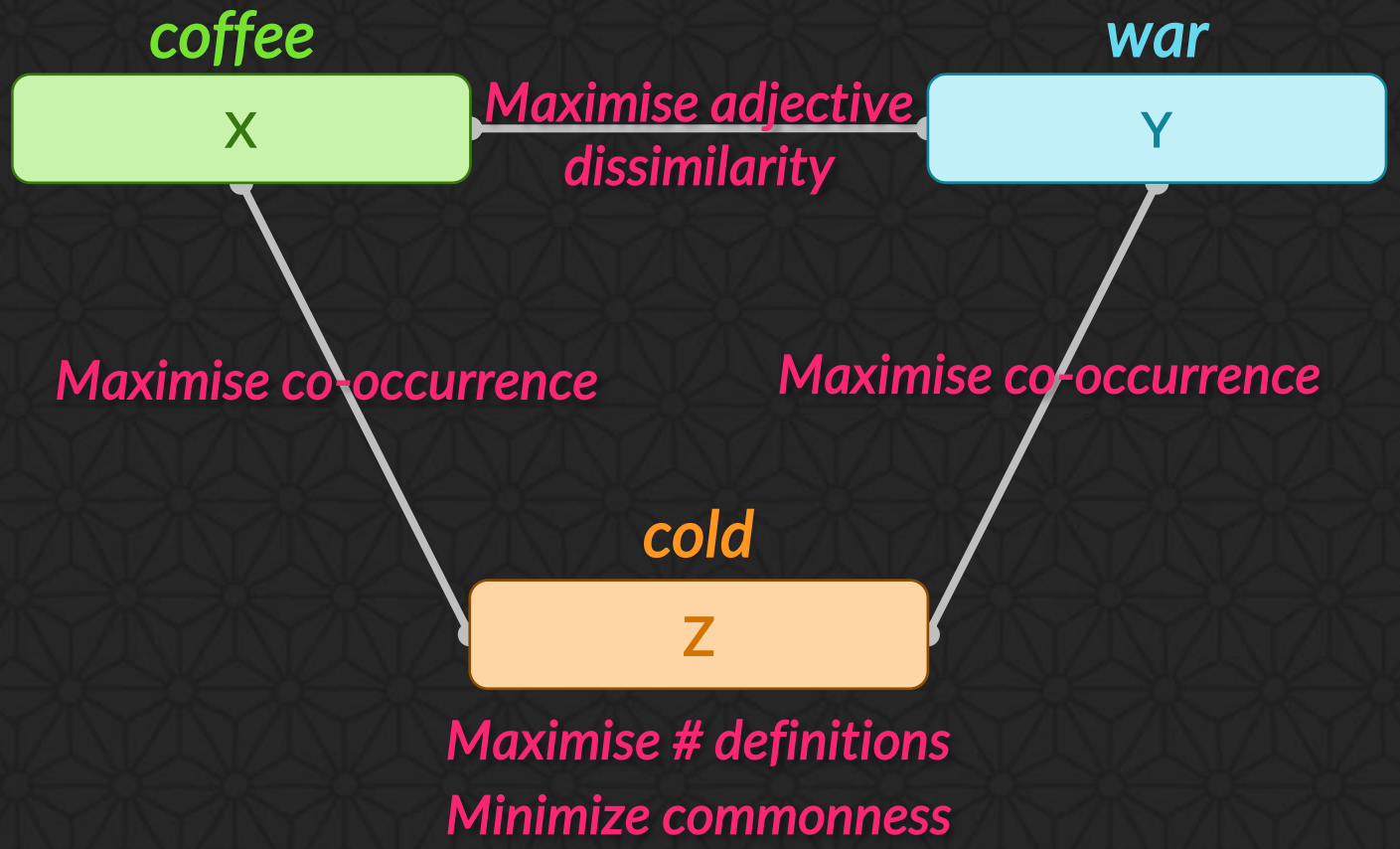
JAPE: Templates & Schemas





Unsupervised Analogy Generator

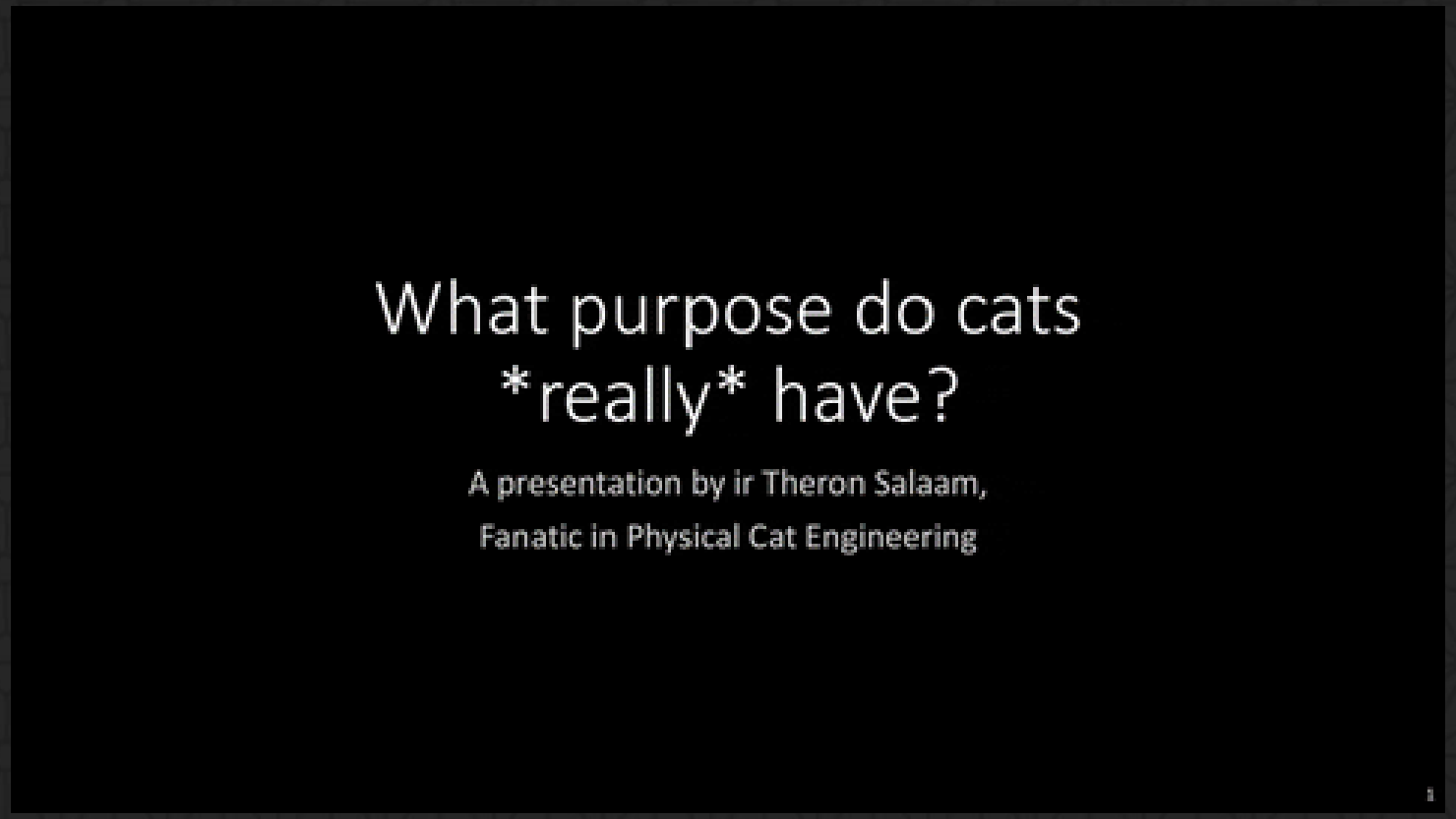




Petrović, S., & Matthews, D. (2013). Unsupervised joke generation from big data

Talk Generator

Generates nonsense PowerPoints about any given topic for presenters to improvise on

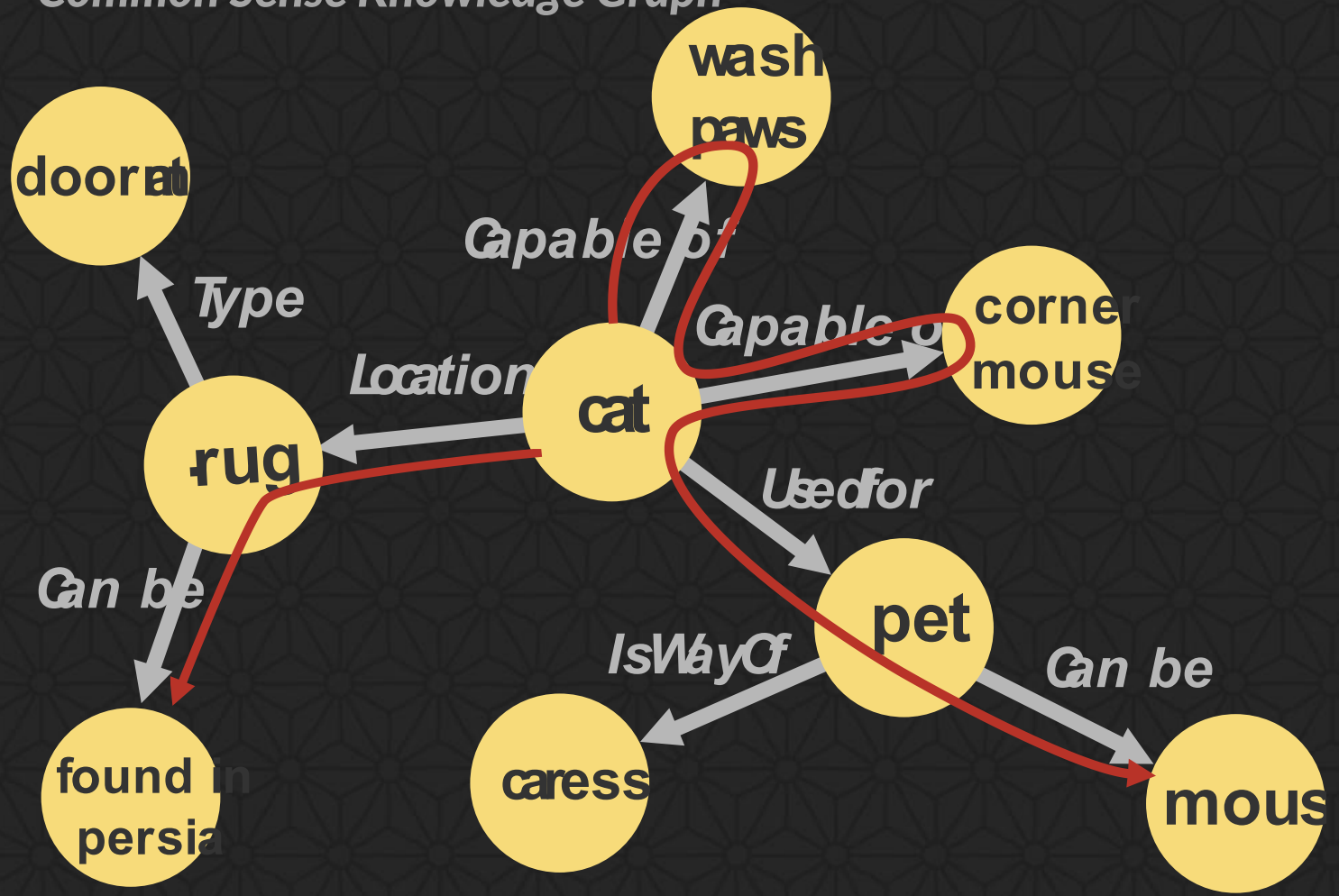


Try it yourself: ***talkgenerator.com***

Slide topic generation

ConceptNet

Common Sense Knowledge Graph



Slide 1	cat
Slide 2	? wash paws
Slide 3	? corner mouse
Slide 4	? pet
Slide 5	? mouse
Slide 6	cat
Slide 7	? rug
Slide 8	? found in persia
Slide 9	cat
Slide 10	cat

Slide generator

[Historical Title Generator]
(Tracery + WikiHow)

[Vintage Person Generator]
(Reddit)

[Quote Generator]
(GoodReads)

[Name Generator]
(Tracery)

corner
mouse

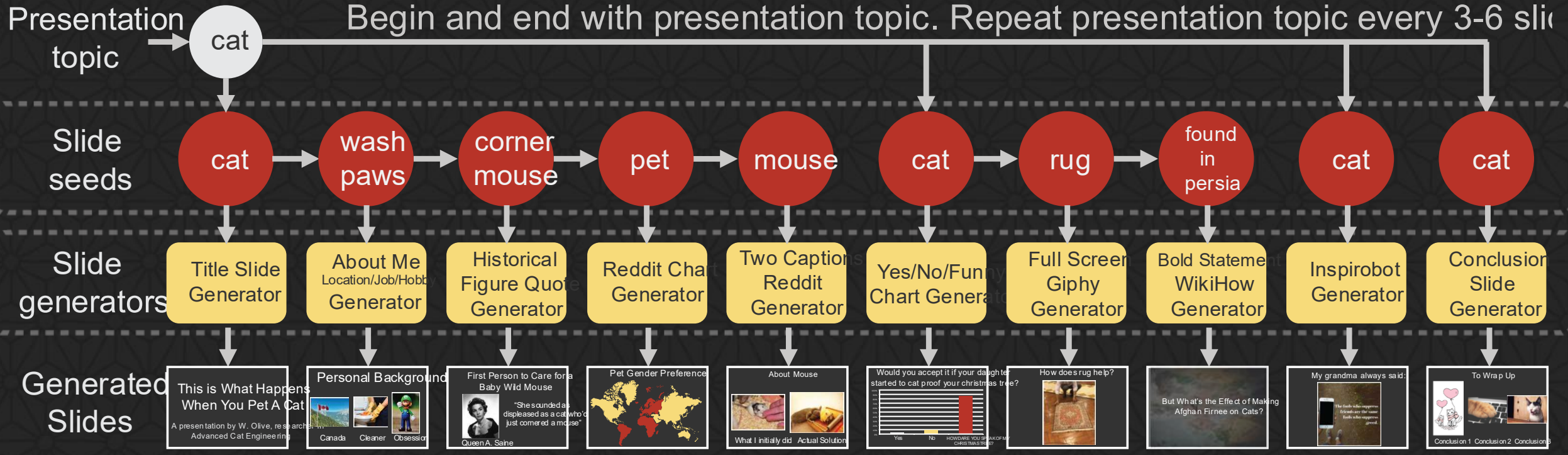
First Person to Care for Baby Wild Mouse



"She sounded as displeased as a cat w just cornered a mou

Queen A. Saine

Talk Generator



Internally uses lots of **handwritten grammars**, combined with **search engines** and **knowledge graphs**

Templates + Grammar + Functions = Babbly

Babbly: my **programming language for text generation** with grammars

```
import firstname.words
food = pasta|pizza|sushi
main = {
  3: <firstname> loves <food.uppercase>!
  1: <firstname> (quite|reasonably|fairly) likes <food>. Oo{1,3}h, I hope they join!
  1: <firstname:protagonist> is not (quite){.5} fond of <food:>.
       <firstname:protagonist> will thus not go to the <food:> (restaurant|place).
}
```

Generates sentences like:

- *Edwin loves PIZZA!*
- *Thomas is not quite fond of sushi. Thomas will thus not go to a sushi place.*

Available on github.com/twinters/babbly
Other popular grammar modelling alternative: ***Tracery***

Used for programming Twitterbot

```
// INPUT
verspreking = "<langwoord.verspreek>"
titel = \titel(<voornaam> <achternaam>)
aanspreking = (<vaste_aanspreking>)/((<titel> /meneer )(<achternaam.verspreek.capitaliseAll>/<voornaam.verspreek.capitaliseAll>)/<voornaam>/Gertje

// SAMSON MANNERISMS

// 'Zo' tussenvoegsel
zo = ( zo){.65}

// Mwah zeg he
ao = (a|o)
mwazeg = Mw<ao:>{1,3}h (zeg hé( <aanspreking>){.6}|<aanspreking>)

// Vraagstelling
vraag = {
    4: wat is dat<zo> een beetje<zo>, een <verspreking><zo>?
    1: wat bedoel je<zo> een beetje<zo> met<zo> een <verspreking><zo>?
}
uitlegVraag = {
    4: <vraag>
    1: ik denk zo stil<zo> in mijn hoofd<zo> dat ik<zo> niet weet wat dat<zo> is<zo>, een <verspreking><zo>.
    1: ik weet<zo> niet<zo> goed<zo> wat dat is<zo>, een <verspreking><zo>.
    1: mag ik<zo> eens<zo> iets<zo> vragen<zo>? <vraag.capitalise>
}

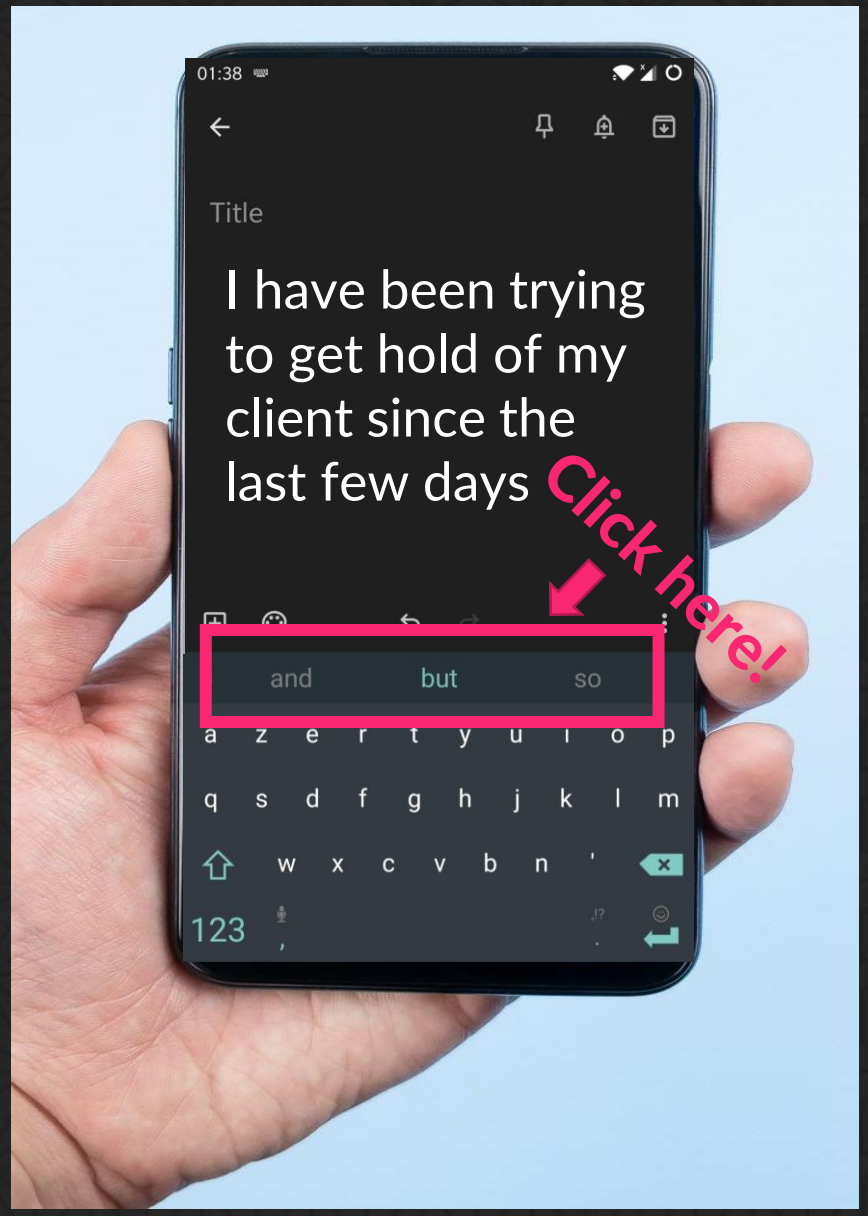
// OUTPUT
main = {
    <mwazeg>, <uitlegVraag>
}
```



Transformer revolution: GPT-3



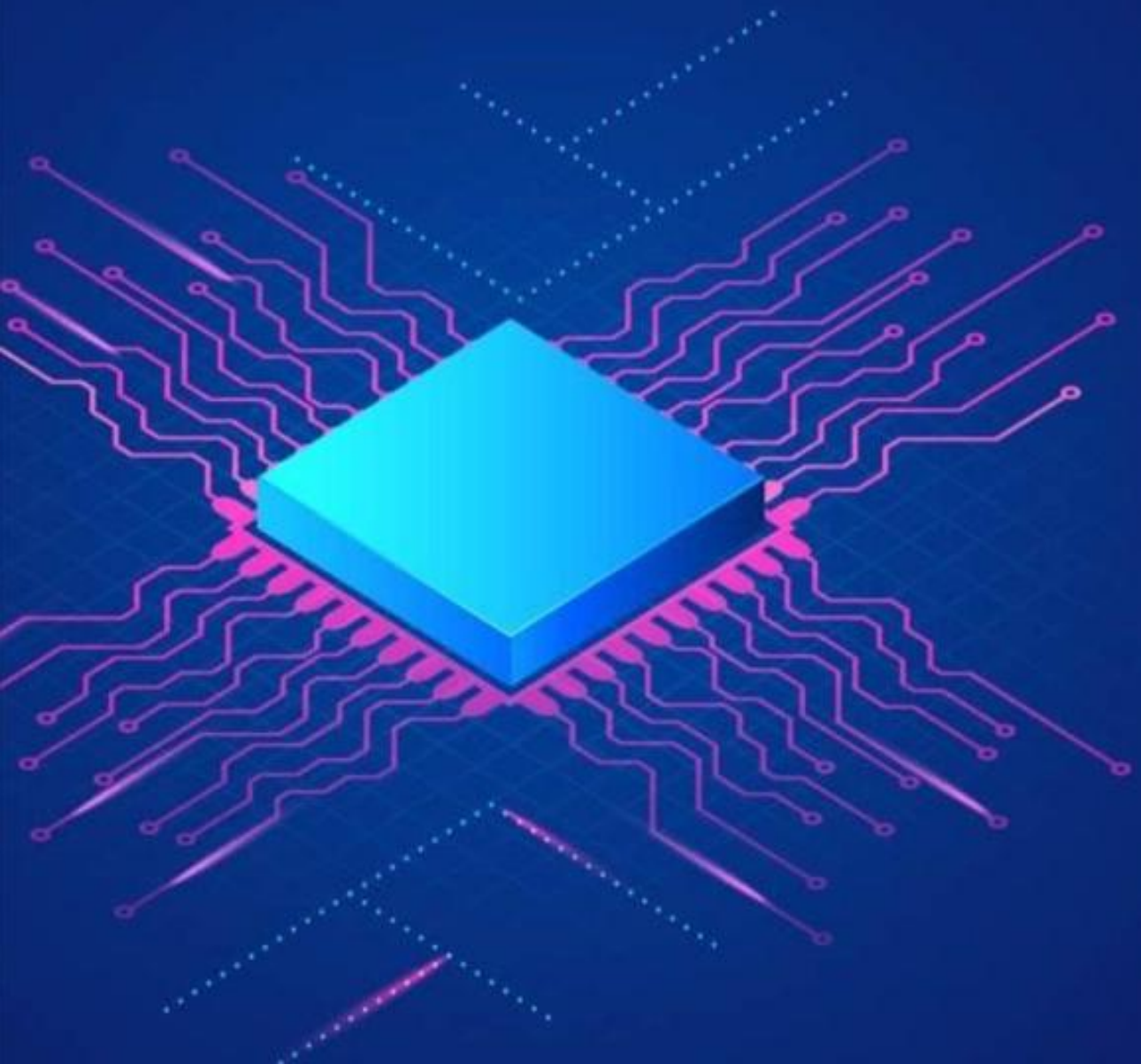
How do statistical text generators work?



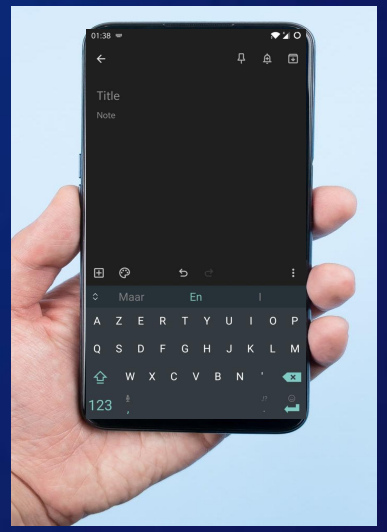
Demo

1. Open your smartphone keyboard *(e.g. in notes)*
2. Press on any suggested **autocomplete** word
3. Press **10-20 times** on a random suggestion
4. You've just generated a sentence using an AI trained to **sound similar to you**! *(if you squint your eyes)*

Autocomplete **counted** how often you used **certain words after other words** in previously typed texts → Statistical model

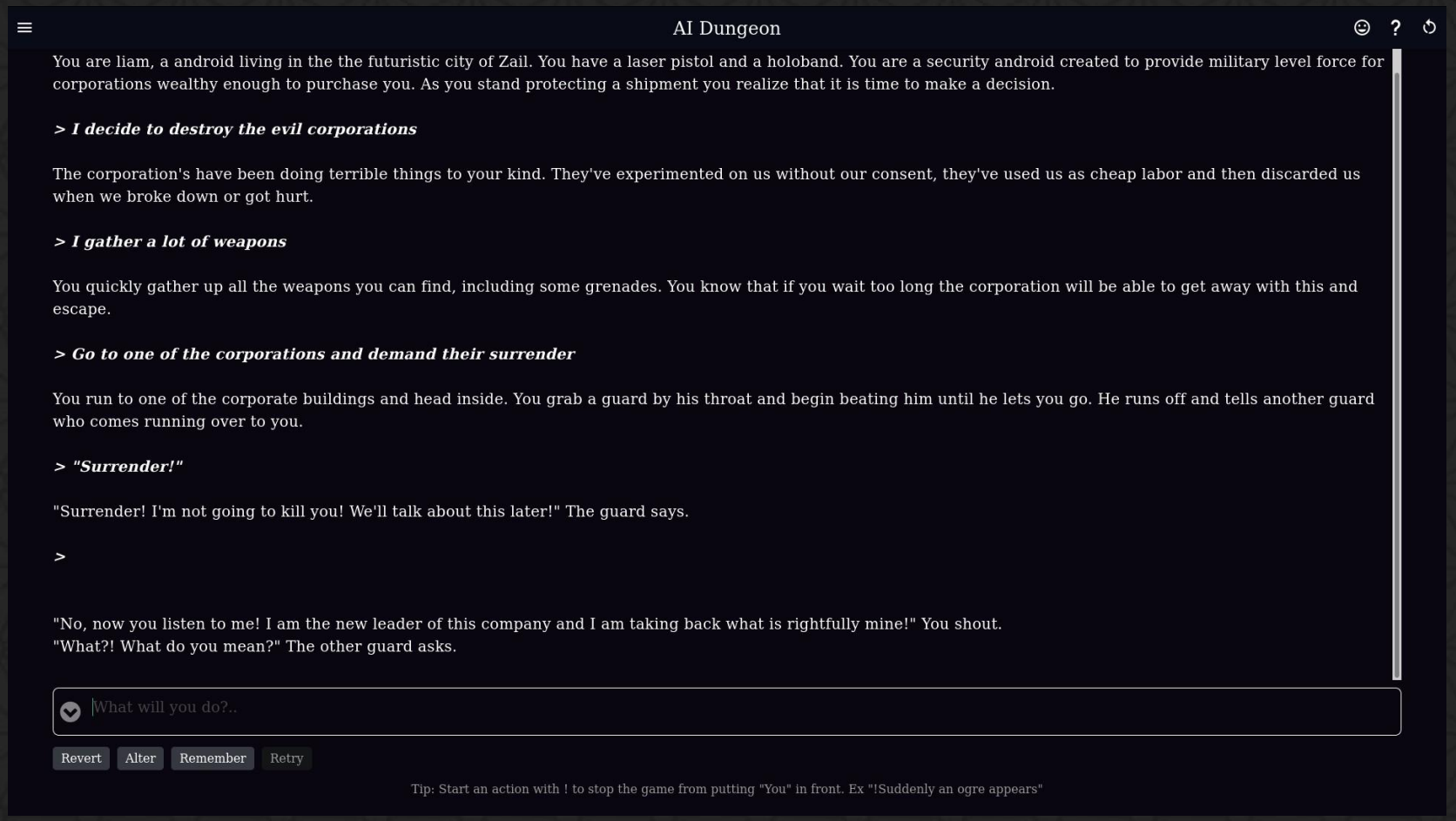


GPT-3



≈ autocomplete on steroids

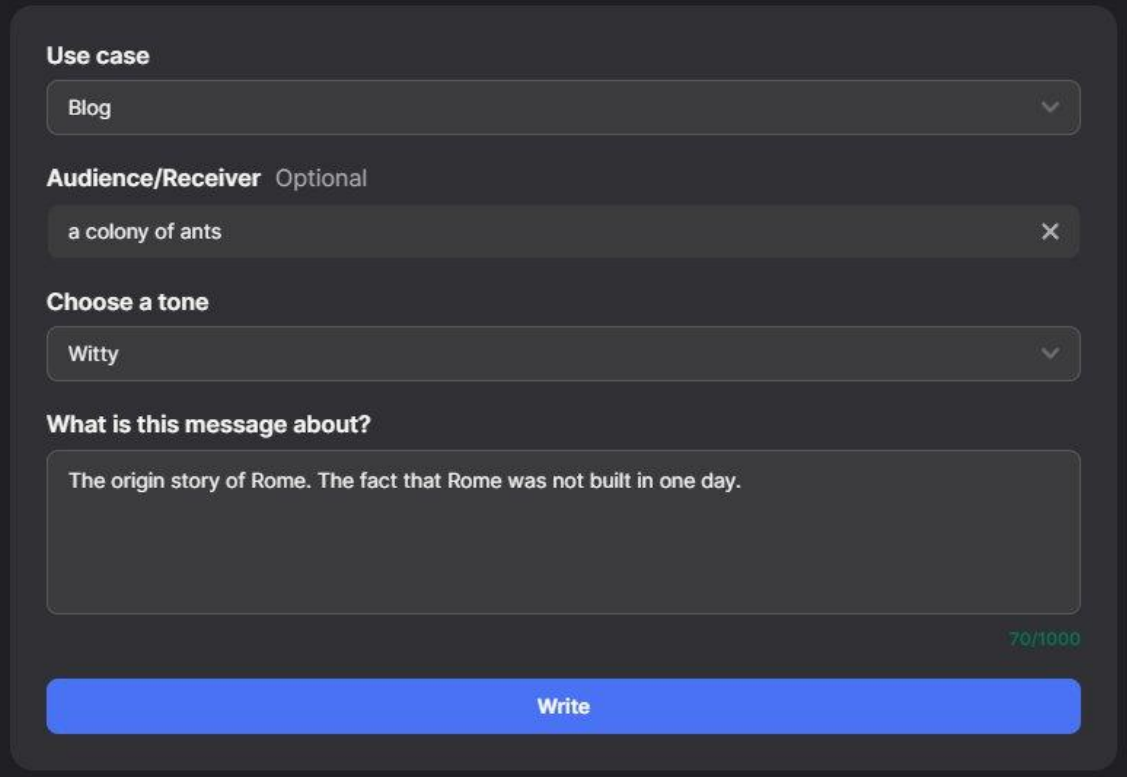
Fun GPT-3 application: AI dungeon

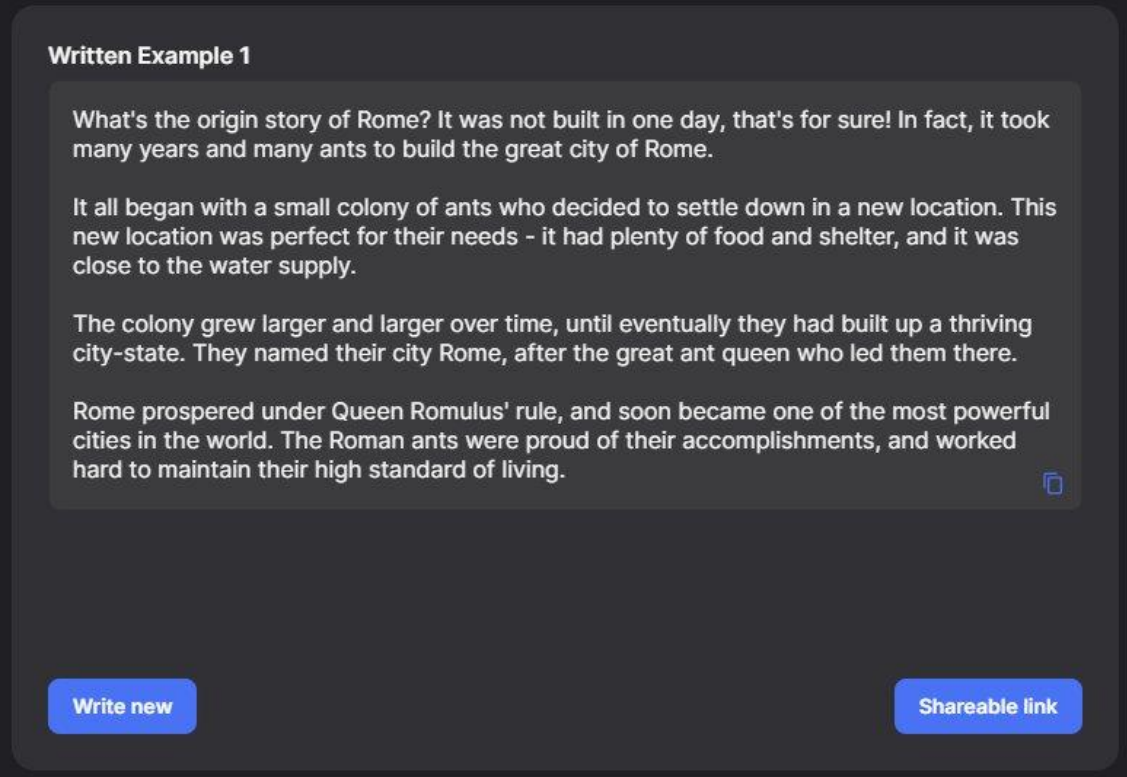


play.aidungeon.io

AI-generated text-based adventures where you can do anything

Fun GPT-3 application: YouWrite

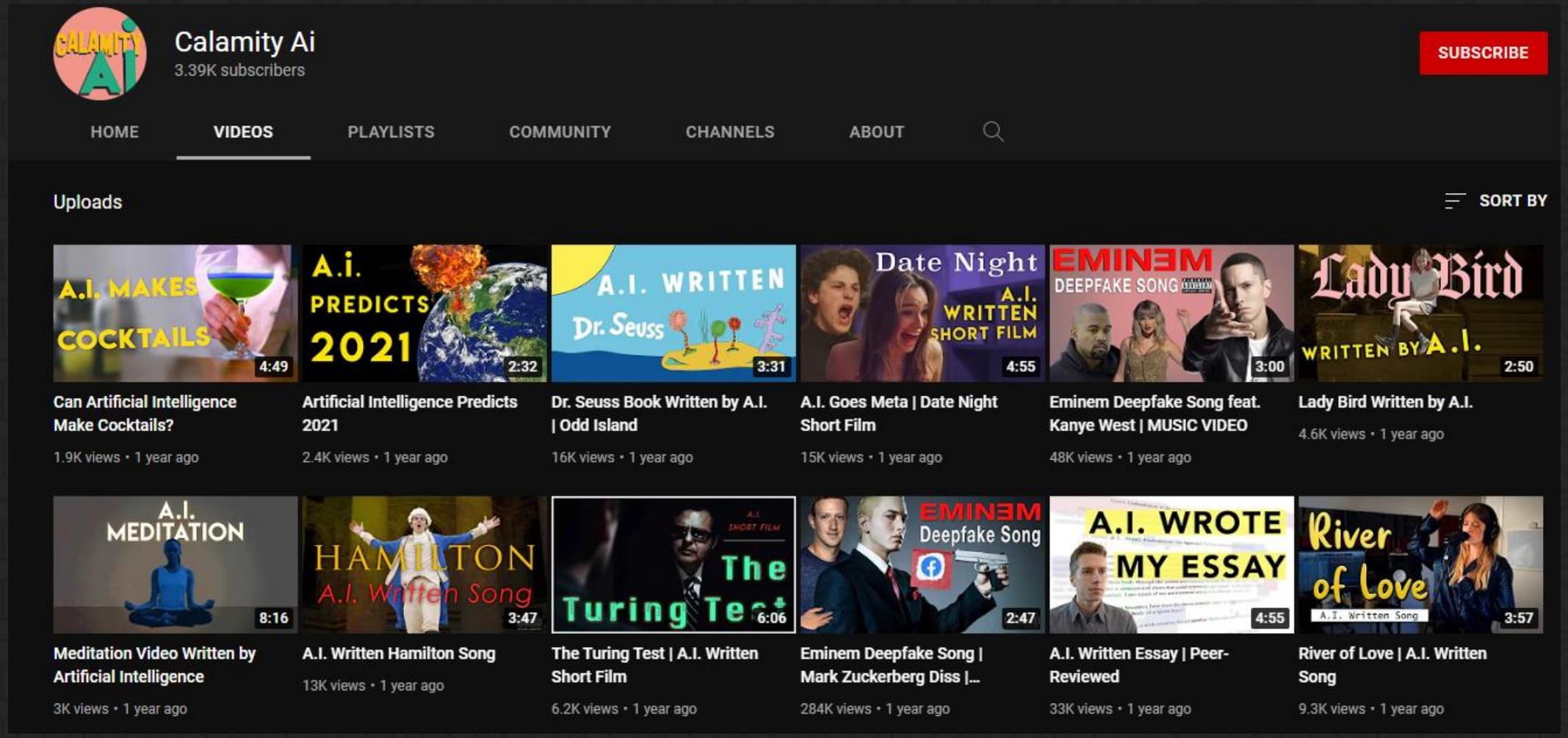




you.com/write

Helps you write any social text

GPT-3 powered theatre script writing



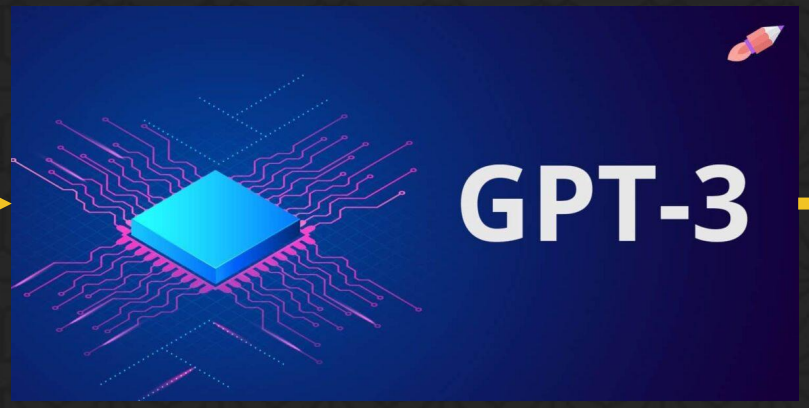
https://www.youtube.com/c/CalamityAi

Getting dialogue out of GPT-3

"-" trick for getting it to generate dialogue

Input

- Fellow ants, we have great news. Our new ant queen is born!
- Hello my loyal ants, I promise I will rule our nest with an iron claw!



Outputs

- All hail our new ant queen!

———————

- I am the new ant queen, and I will make our colony the strongest in the world!

———————

- It's time for us to expand our nest!
- Our nest is under attack!



Improbotics



A.L.Ex the robot **plays improv scenes** with improv actors

“Listens” to scene and **formulates own responses** with GPT-3

Later in show: gains **human body**. AI whispers through ear piece what actor should say.

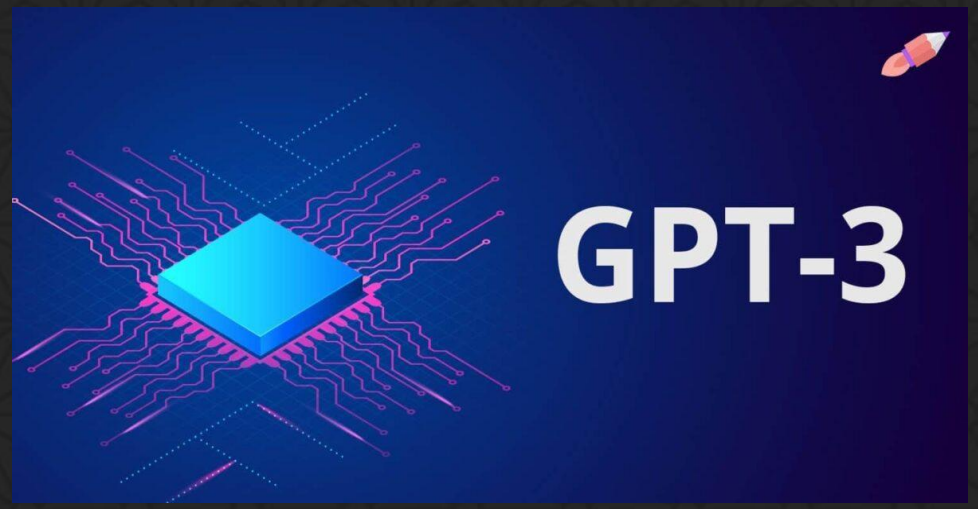
End: everyone has ear piece
→ **Turing test**!

Improbotics interface



Operator types what happens and selects responses

Access



You can get GPT-3 access by signing up here: https://openai.com/api/

HUGGING FACE

Or if you like smaller models that you can run for free on your own pc *(e.g. GPT-2 or other transformer alternatives)*: https://huggingface.co/models

Learning a sense of humor?

Transformer models

Large **language models**, **pretrained** on large corpora





GPT-2 & GPT-3

Completes any textual prompt

BERT

Classifies any text sequence / token

Outperforming previous neural architectures
on **most language tasks**

Brown, Tom B., et al. "Language models are few-shot learners."
Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding

Not just for English!



RobBERT

Our Dutch BERT-like model

Phase 1: Pretraining:

- Find unlabeled **corpus** for target language
- **Expensive** pretraining phase
- BUT: **reusable**! Just download existing!

Phase 2: Finetuning a pre-trained BERT

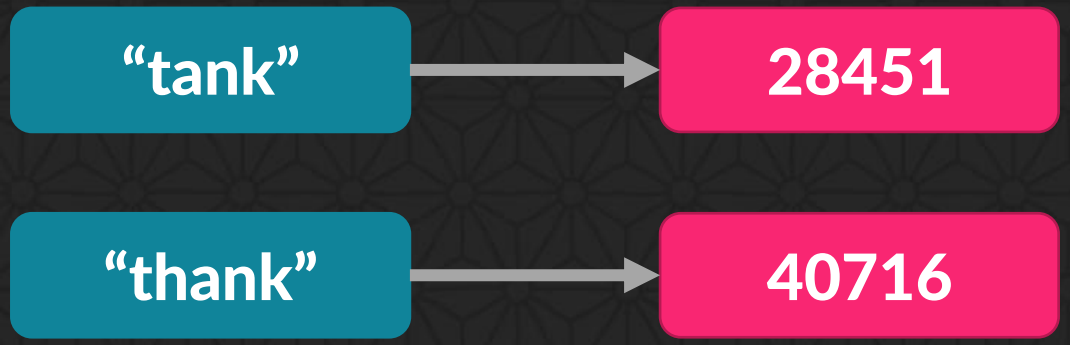
- Feed **labeled** training data
- Relatively **cheap** fine-tuning for the target task
- Usually **outperforms** other language models on most NLP tasks

Delobelle, P., Winters, T., & Berendt, B. (2020). RobBERT: a dutch RoBERTa-based language model.

GPT-2 / GPT-3 humor?

Tokenizer problem: maps words to numbers

→ **Remove** knowledge about **letters**!



Hard to make/recognize new **wordplay**!

https://www.gwern.net/GPT-3

https://towardsdatascience.com/teaching-gpt-2-a-sense-of-humor-fine-tuning-large-transformer-models-on-a-single-gpu-in-pytorch-59e8cec40912

GPT-2 / GPT-3 joke examples

GPT-2 can mimic joke style, but usually very absurd

What did one plate say to the other plate?
Dip me!

What did the chicken say after he got hit by a bus?
"I'm gonna be fine!"

A woman walks into the bar.....she says to the bartender "I want a double entendre" and the bartender says "No, we don't serve that"



https://www.gwern.net/GPT-3
https://towardsdatascience.com/teaching-gpt-2-a-sense-of-humor-fine-tuning-large-transformer-models-on-a-single-gpu-in-pytorch-59e8cec40912

Reminiscent of how children write jokes





Let's try simpler & more understandable models!

Dynamic Templates

"Are there also atheists that don't believe in atheism?"

Base text words:

b_0 b_1 b_2 ... b_n

POS tag

b_0	noun,pl
...	...
b_n	verb,1pl

Content text words:

c_0 c_1 c_2 ... c_n

POS tag

verb,1pl	c_0
...	...
adj	b_n

"They see the fact that the former God is not trying to deny this newly acquired insight as proof of them being right. Even with the Church, things are not going well. Norse popes."

Replace some matching words with matching POS tag with lowest frequencies

b_0 c_2 b_2 ... c_0

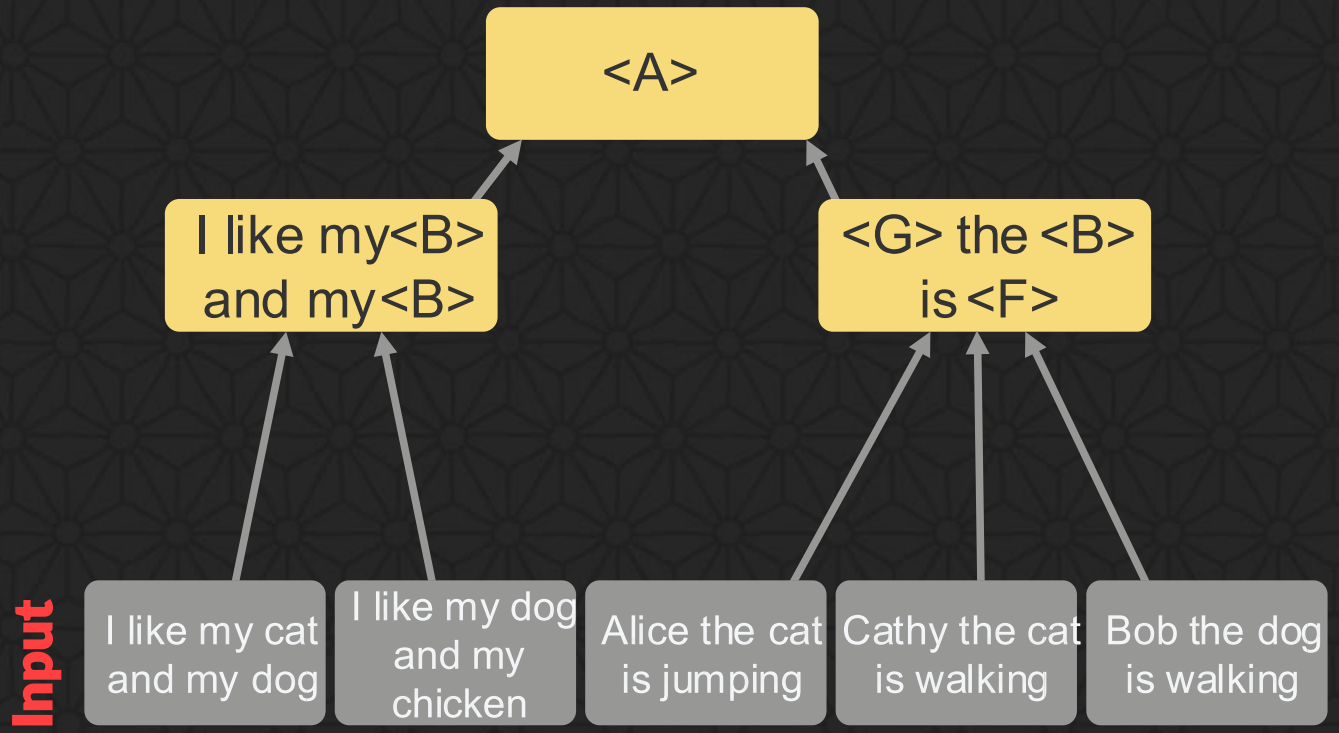
"Are there also popes that don't believe in God?"

Used in @TorfsBot



GITTA: Template Trees for extracting templates

1. Join **closest** strings
2. **Merge similar** template slots
3. Iteratively **simplify until convergence**



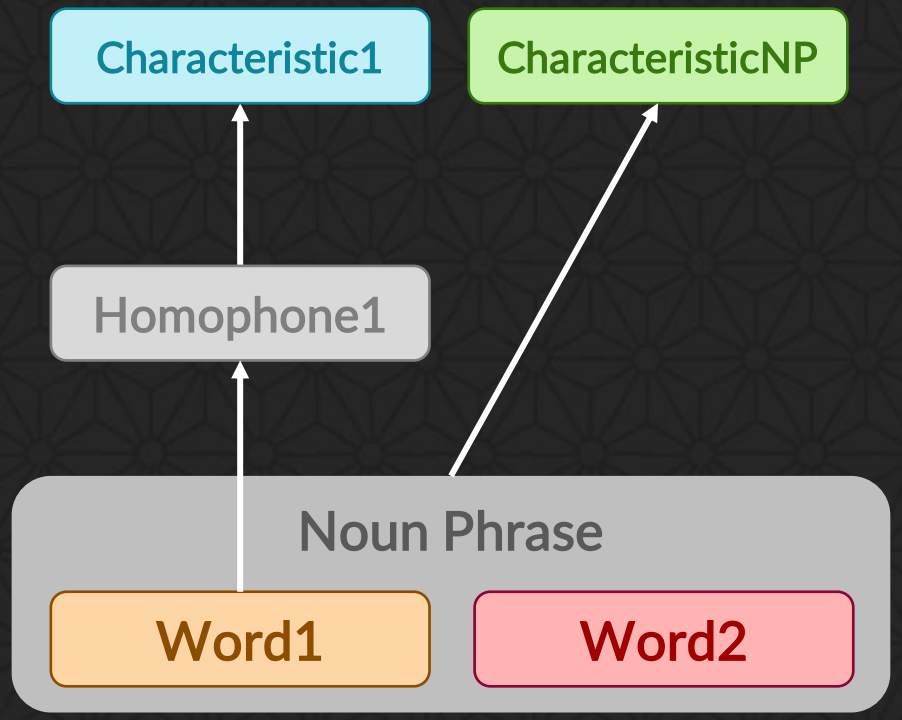
A: I like my <B> and my <B>
| <G> the <B> is <F>

B: chicken | cat | dog

F: walking | jumping

G: Alice | Bob | Cathy

Generalising Schemas to Enable Machine Learning



Constraint-based Schemas

Hard to tweak to preference

Header

Keywords

Template

Features





Aggregator

Values Generator



Metrical Schema

Allows for learning in aggregator component

GOOFER: Generalised framework

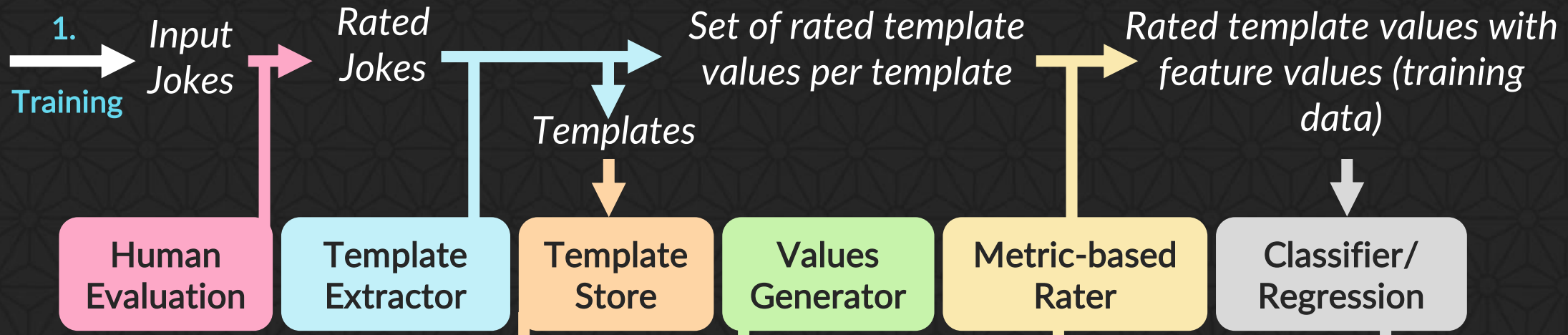
a) I like my relation… n.
b) I like my coffee li…
c) The sailor bears …
d) I like my girlfrien…
e) The lord cultivat…

Joke	Rater 1	Rater 2	Rater 3
a	5	4	5
b	4	4	3
c	5	3	4
d	2	1	3
e	3	2	2

Template 1: "I like my X like I like my Y, Z."
relations, source, open | [5,4,5]
coffee, war, cold | [4,4,3]
girlfriend, estate, real | [2,1,3]

Template 2: ``The A B a C. D.''
sailor, bears, stress, Pier pressure | [5,3,4]
lord, cultivates, region, A baron land | [3,2,2]

X	Y	Z	Score	FreqX	FreqY	FreqZ	FreqXY	...
relations	source	open	[5,4,5]	3 831 210	4 050 904	4884757	0	...
coffee	war	cold	[4,4,3]	784 065	9 211 826	2 010 173	0	...
girlfriend	estate	real	[2,1,3]	75 392	998 669	5 284 057	0	....



"Men"

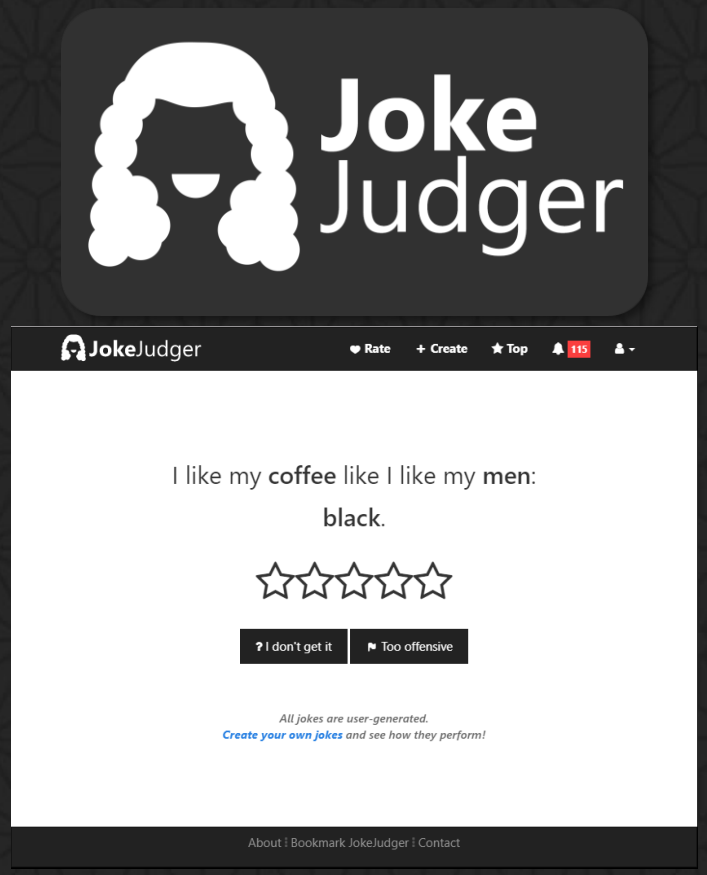
"Men" + *"I like my X like I like my Y, Z."*

a) men, grave, nameless
b) men, country, cold
c) men, turkey, roast
d) men, rain, gentle
e) men, laugh, silly

X	Y	Z	Score	FreqX	FreqY	FreqZ	FreqXY	...
men	grave	nameless	?	9 876 543	4 050 904	4884757	0	...
men	country	cold	?	9 876 543	9 211 826	2 010 173	0	...
men	turkey	roast	?	9 876 543	998 669	5 284 057	0	...
men	rain	gentle	?	9 876 543	9 876 543	845 321	0	...
men	laugh	silly	?	9 876 543	954 569	123 456	0	....

I like my men like I like my grave: nameless
I like my men like I like my rain: gentle

GAG: implements GOOFER framework

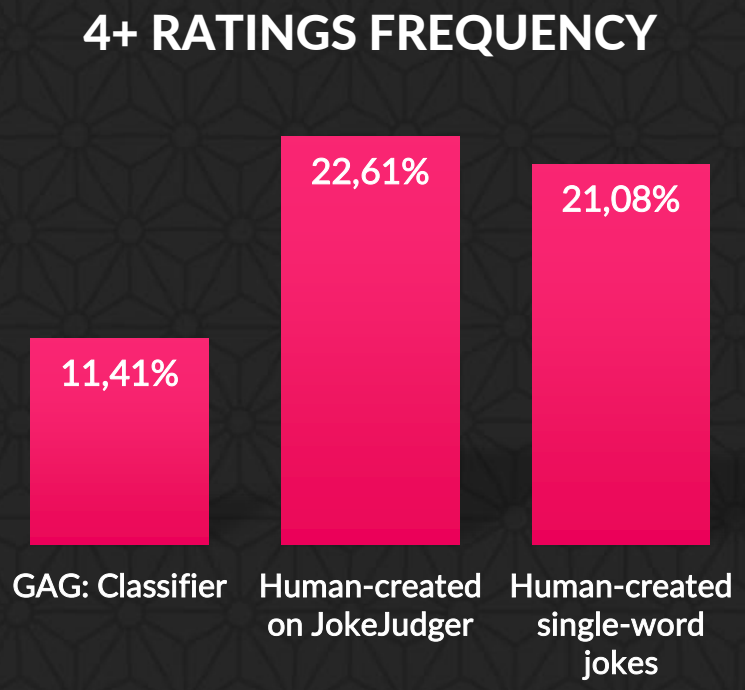


GAG
GENERALISED ANALOGY GENERATOR

I like my **coffee** like I like my **sleep**: **strong**.

I like my **men** like I like my **graves**: **nameless**.

I like my **women** like I like my **tests**: **independent**.



Winters, T., Nys, V., & De Schreye, D. (2018). Automatic Joke Generation: Learning Humor from Examples

Need for good humor detection algorithms!

Part 3: Humor Detection

Early Humor Detector

- Designed **humor features** *e.g. alliteration, antonym, adult slang…*
- Used **Naive Bayes** and **Support Vector Machines**
- **Task:** One-liners vs **news**, **neutral corpus** & **proverbs**

One-liners Reuters	One-liners BNC	One-liners Proverbs
96.95%	79.15%	84.82%

Table 4: Humor-recognition accuracy for combined learning based on stylistic and content features.

Mihalcea, R., & Strapparava, C. (2005). Making computers laugh: Investigations in automatic humor recognition.

But is this a good dataset?

News & proverbs have **completely different types** of words than jokes!



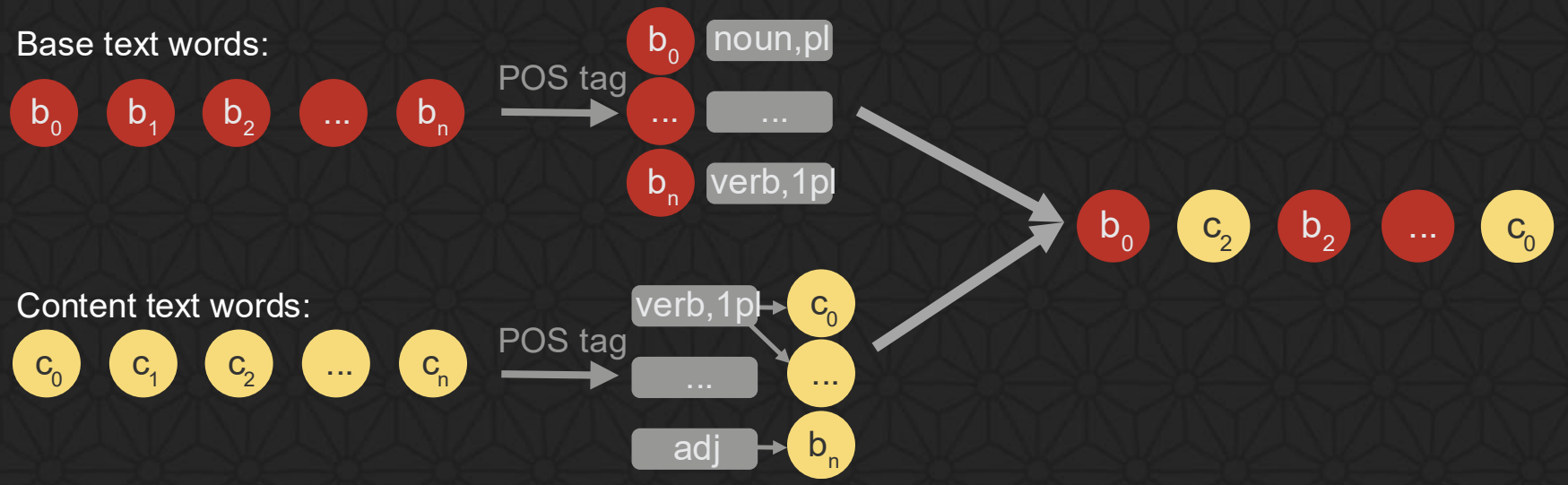


→ Looking at **word frequencies** is often already *"enough"*!

Is this really **humor** detection?

Jokes are **fragile**!

Two men are in a bar. Says one to the other:
"Do you know how to drive this thing?"



Generate non-jokes using **dynamic templates**!

Word-based features **won't work** anymore!

Examples of generated Dutch non-jokes

Het is groen en het is een mummie?
Kermit de Waterkant

Wat is het toppunt van principe?
1) Wachten totdat een Nederlander gaat twijfelen
2) Een Zuster met een autoladder
3) Een brandwacht brandmeester met een brandmeester van 9 maanden

“Ober, kunt u die schrik uit mijn politieman halen? Want ik eet liever alleen.”

"Mijn hond is heel vreselijk: Hij schreeuwt mij iedere zus de broer.“
"Maar dat is toch niet zo heel vreselijk?“
"Jawel, want ik heb geen rapport!"

Wat staat er midden in het bos?
De kapper.

Er loopt een super vriendelijk blondje langs een armband. Last er een toonbank: “zo, waargaan die mooie mannen heen?” Blondje: “naar de barkeeper als er niets tussen komt…”

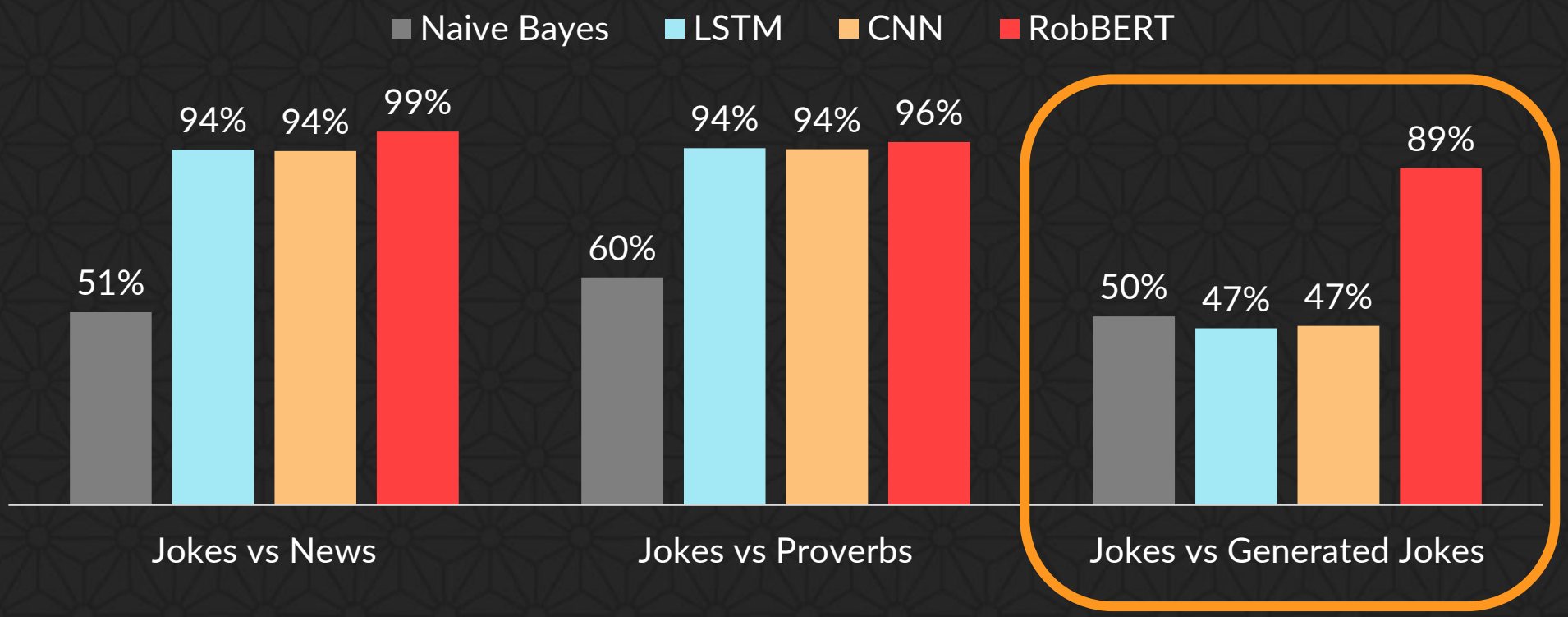
Hoe heet de vrouw van Sinterklaas?
Keukentafel.

"Twee tanden zwemmen in de zee en ze zien een stamgast op een stamgast. De ene raad zegt tegen de andere raad: 'Hé kijk! Ons eten op een bord!'"



Much more challenging dataset!

Binary classification of jokes versus texts from other domains



More truthful humor detection?

Two RoBERTa models with different heads

Hey, I'll just alter the headline one word at a time!



Alright! I'll check if that edit actually makes it funnier!

Satire MLM

Finetuned RoBERTa MLM on all headlines and satire

Sentence with mask token → suitable token filled into mask

Dog <mask> bites man → Dog accidentally bites man

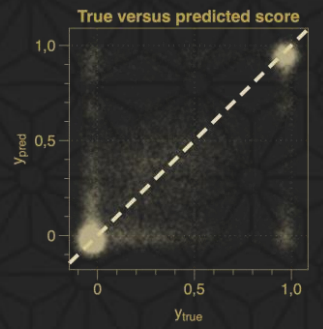
→Used for mutation operator: substituting & adding random tokens

Satire Regression

Finetuned RoBERTa regression head: predicts headline (0), satire (1) or on perceived funniness in between ([0,1])

"Most Americans Want Congress To Investigate Michael Flynn"
→ Predicted score: 0.01

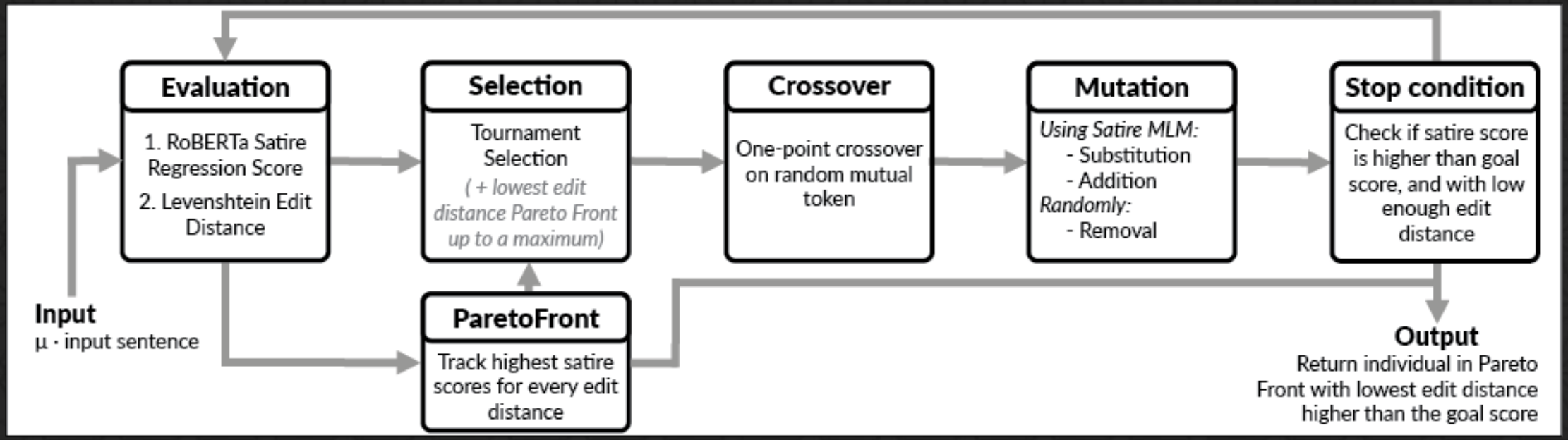
News: The Perils Of Climate Change: A Rise In Sea Level Of Just One Foot Would Be Devastating For The Man Tied To This Pier
→ Predicted score: 1.05



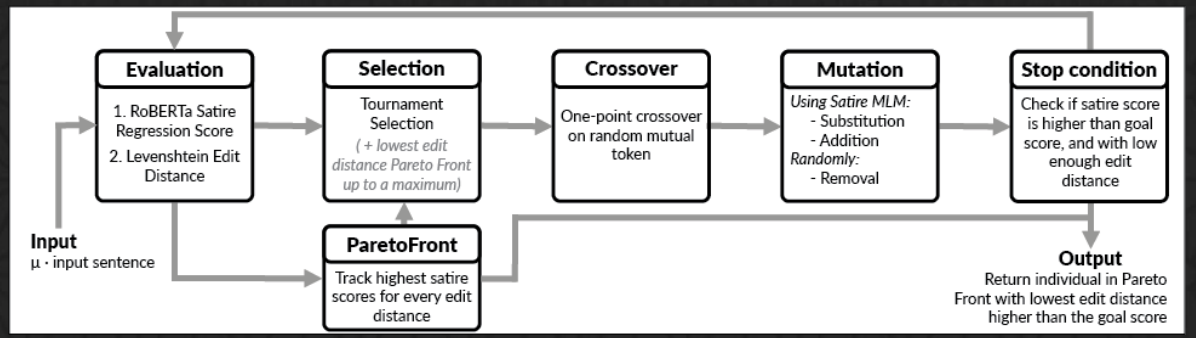
→ Used for fitness function

GALMET: Genetic Algorithm using Language Models for Evolving Text

Created algorithm that evolves headlines into satirical headlines



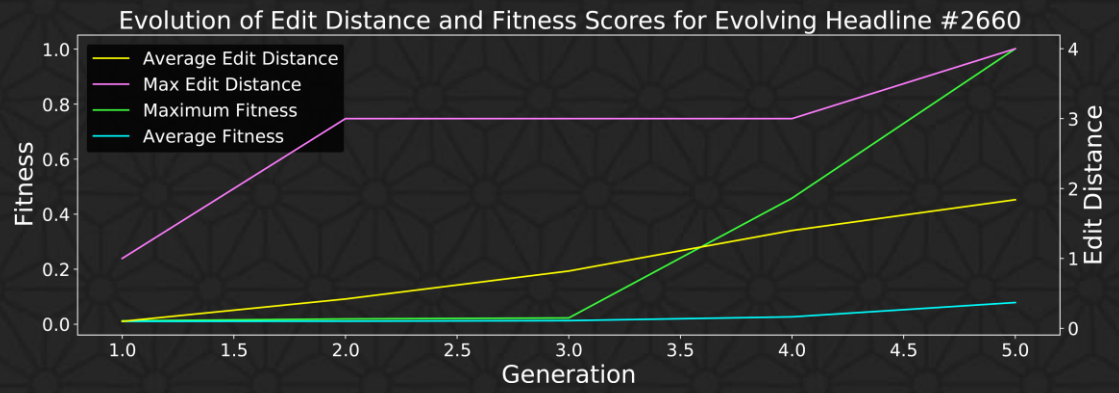
Tiger bites zoo visitor



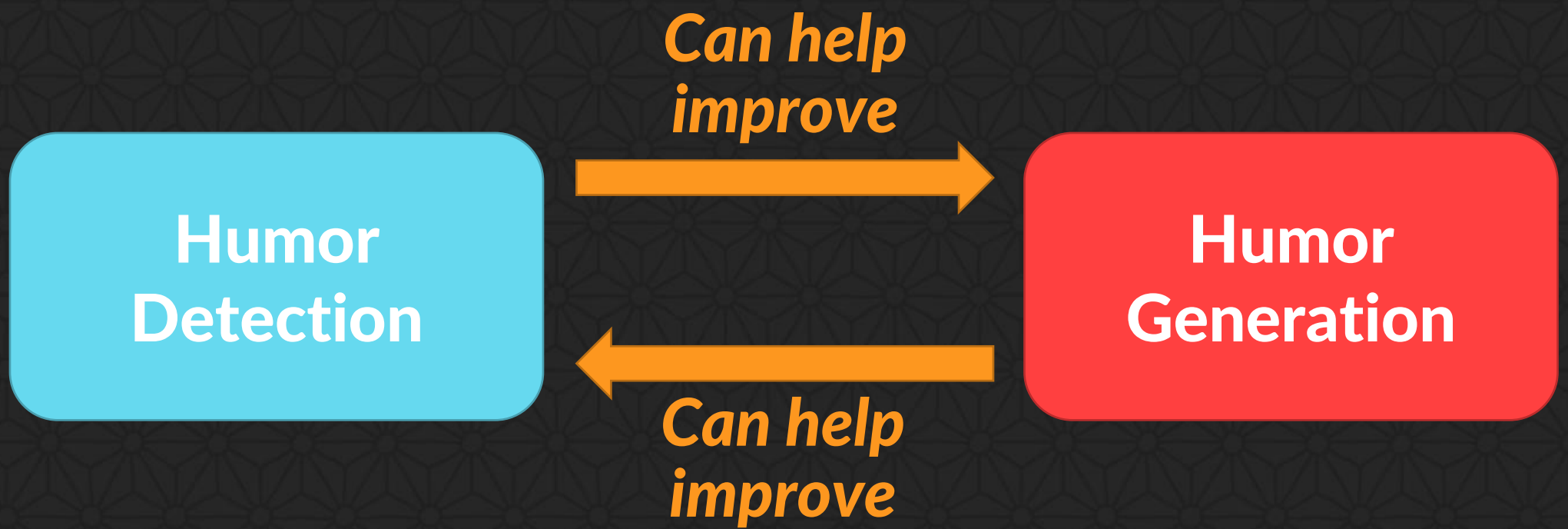
Survival of the Wittiest

Most Americans Want Congress To Investigate Michael Flynn
→ 224 Americans Asked To Investigate Michael Jordan

d_{edit}	Score	Phenotype
0	0.010341	Most Americans Want Congress To Investigate Michael Flynn
1	0.019262	Why Americans Want Congress To Investigate Michael Flynn
2	0.070894	Most Americans Want To Investigate Michael Jordan
3	0.458300	224 Americans Asked To Investigate Michael Flynn
4	1.001400	224 Americans Asked To Investigate Michael Jordan

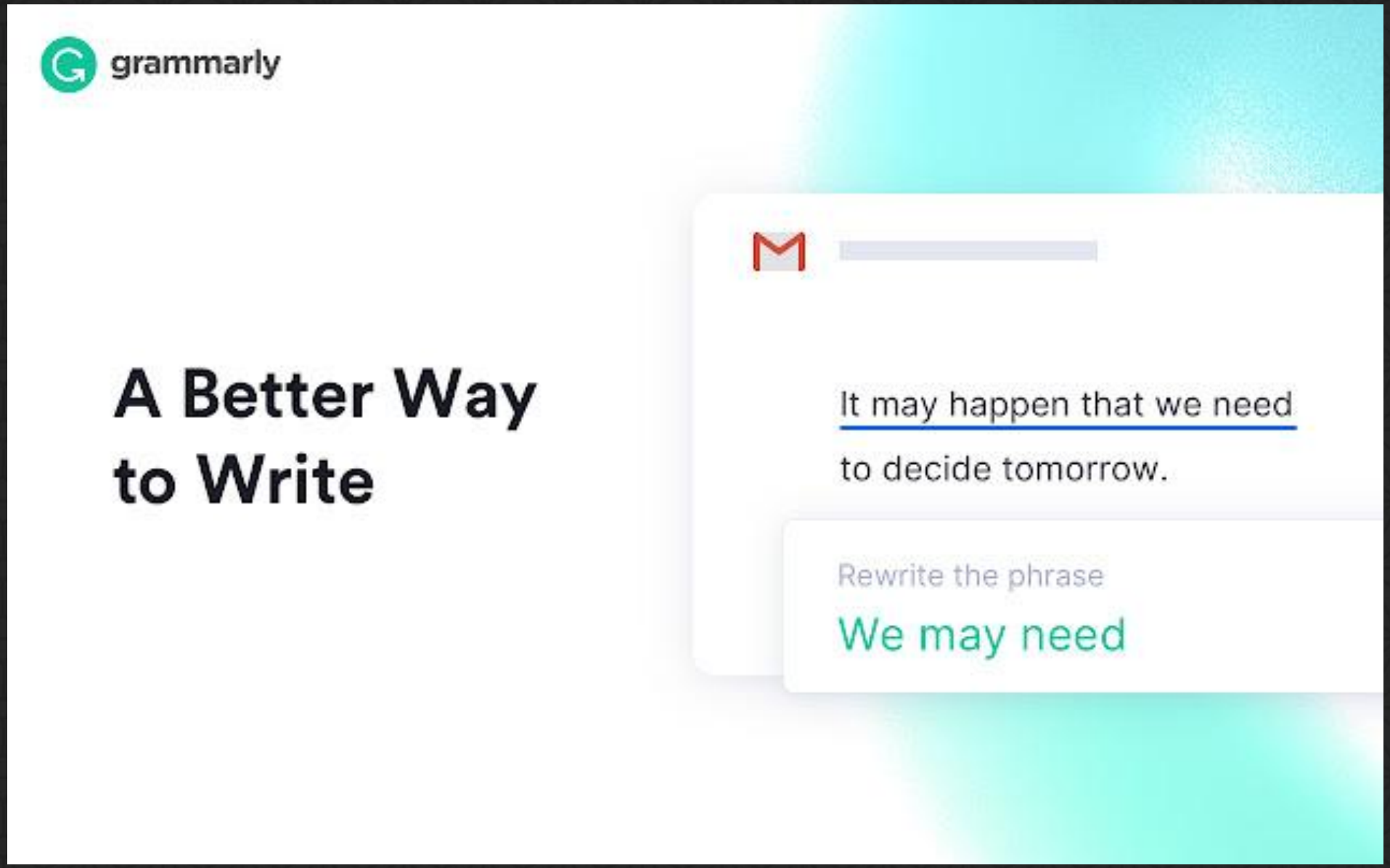


Symbiotic relationship



Part 4: Why?

Grammarly but for jokes?



VIRTUAL ASSISTANTS

People often treat virtual assistants like humans

("please", "thank you", "sorry") [1]

41%

say it's like talking to a friend [1]

Humor and creativity in language is important **between friends** [2]



[1] https://www.thinkwithgoogle.com/consumer-insights/voice-assistance-consumer-experience/
[2] Laughter's Influence on the Intimacy of Self-Disclosure, Gray, A.W., Parkinson, B. & Dunbar, R.I. Hum Nat (2015) 26: 28.

Casual Creation: it's fun to create these little bots!



https://thomaswinters.be/mopjesbot

Mutually interactive bots



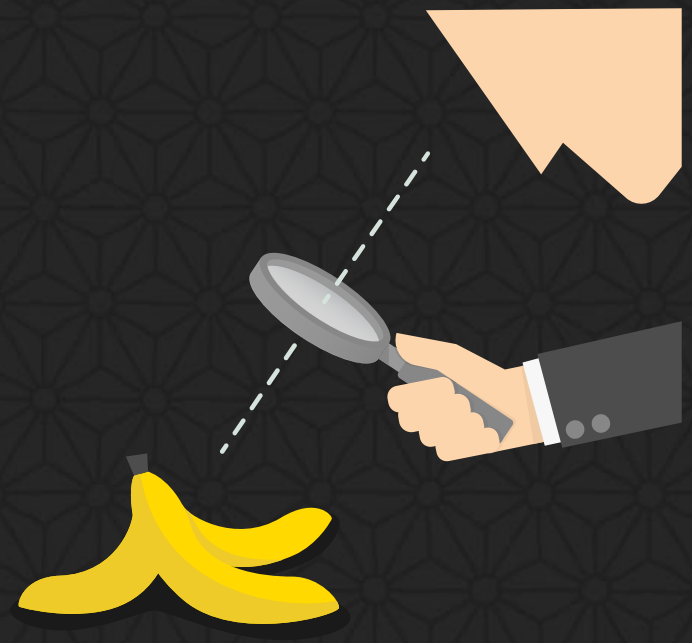


https://thomaswinters.be/samsonbots

Studying & understanding humor

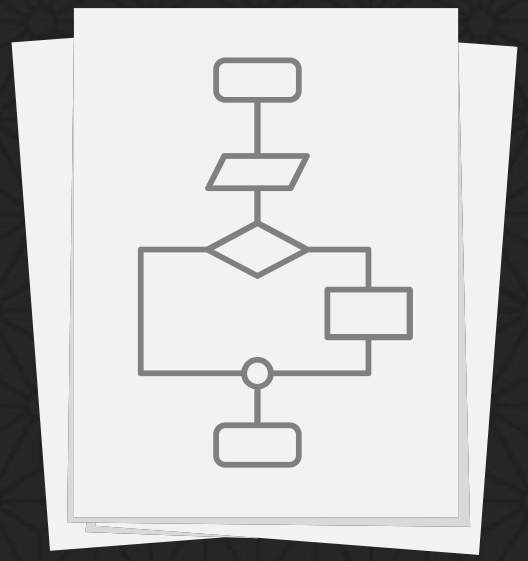
We do not fully know how humor works



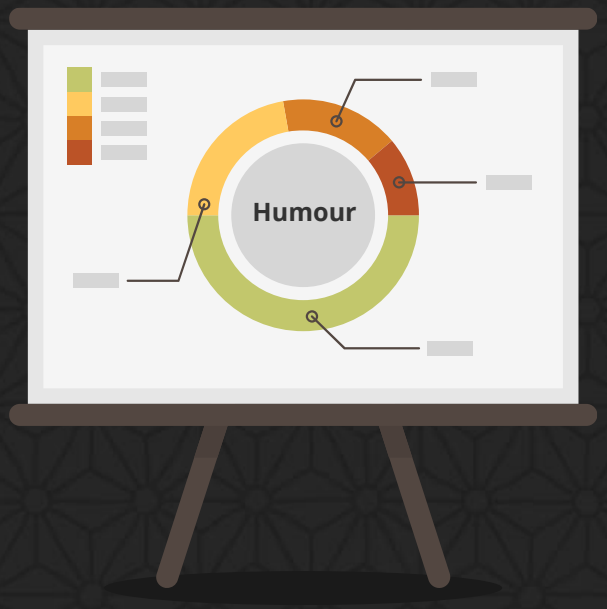


Artificial Intelligence can shed a light!

Can computers have a sense of humor?



Hand-craft rules
for type of humor



Learn rules if
enough data



Better **NLP models**
get us closer & closer

BUT:

AI-complete problem!

Need to process two mental images
before fully understanding the joke

Computational humor is still in its infancy



But, children have a fascinating sense of humor



Thank you!

Thomas Winters

PhD Student at KU Leuven & FWO Fellow

@thomas_wint

thomaswinters.be

KU LEUVEN DTAI fwo

Some images (based on the works) of dooder on freepik.com