Hoe leer je computers humor?



Thomas Winters

PhD Student at KU Leuven & FWO Fellow

@thomas_wint

thomaswinters.be

What is the difference between leaves and a car?





One you brush and rake,

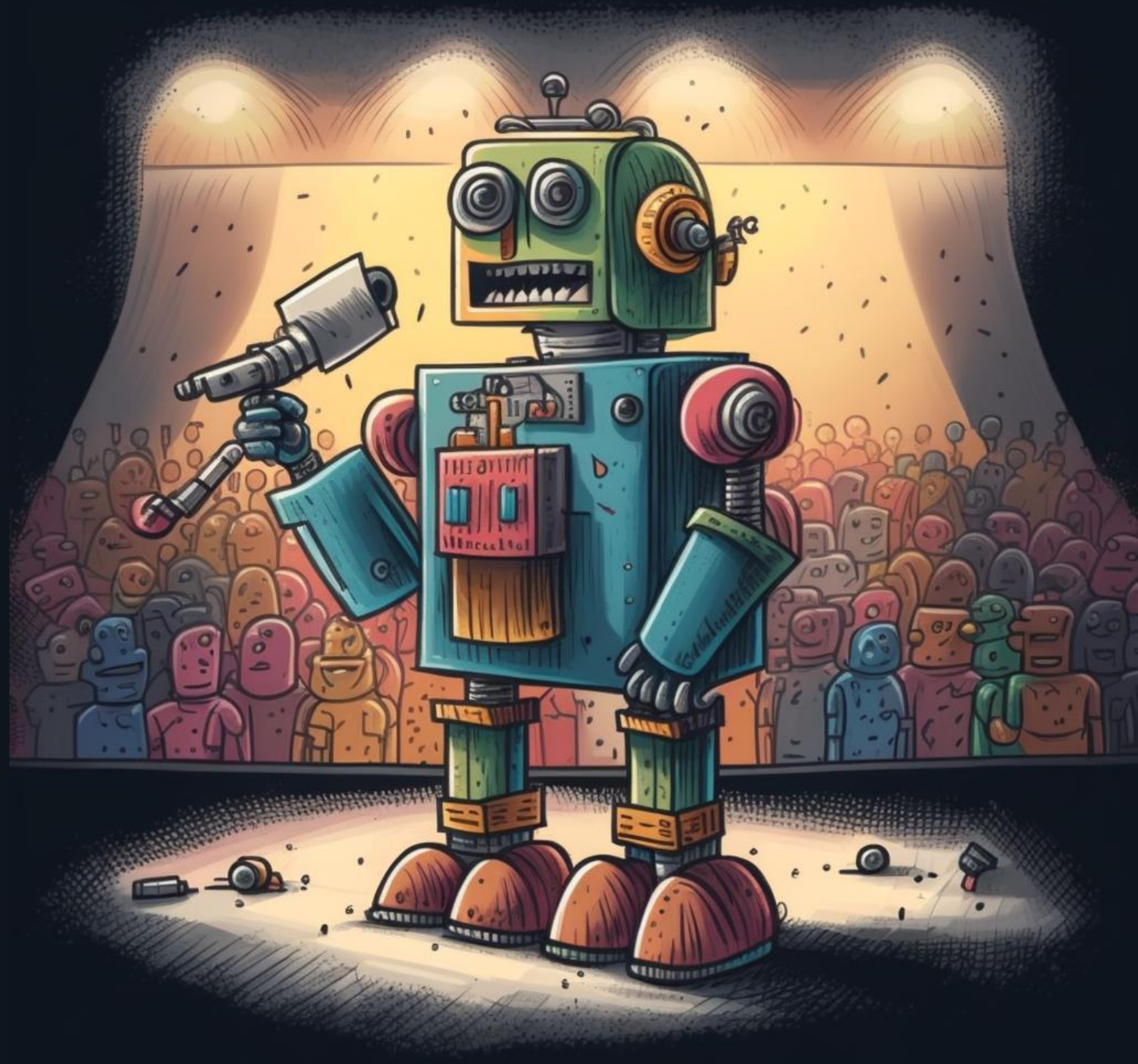
the other you rush and brake.



Wat is het verschil is tussen een gewoon festival en het Nerdland Festival?

Als je op het Nerdland Festival alcohol wil krijgen, moet je je bestelling afwegen in mol!

Computer + humor = paradox?



Humor = intrinsiek menselijk



Waarom hebben computers humor nodig?





Reden 1:
Verlagen van de humor leercurve

Humor co-creatie

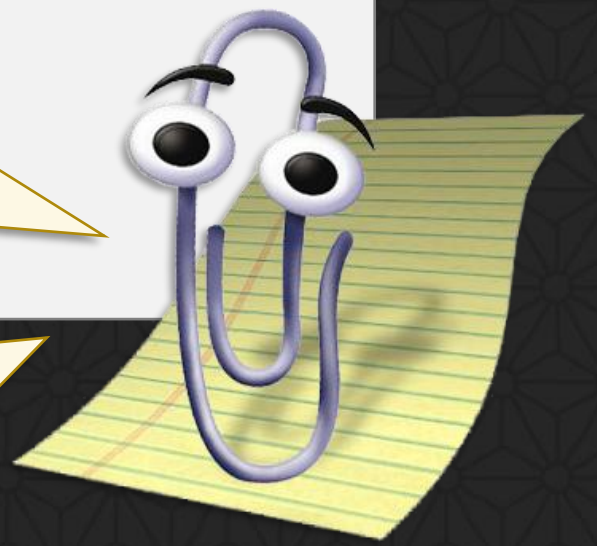
Grammarly / Clippy voor mopjes

Mijn ex mist me nog steeds, maar ze mikt steeds beter

En toen hebben ze een duif gebruikt om een bom op een schip te laten vallen

Specifieker zijn kan helpen bij grappigheid, bv.
“ingezet om een explosief op een nietsvermoedend schip te laten ploffen”

Grappige bedenking
“ironisch om het teken van vrede te gebruiken als oorlogswapen”



Reden 2: AI assistenten moeten herkennen wanneer je grapt



Mensen behandelen assistenten al menselijk

(“alsjeblieft”, “dankjewel”) [1]

Humor is belangrijk in vriendschap [2]

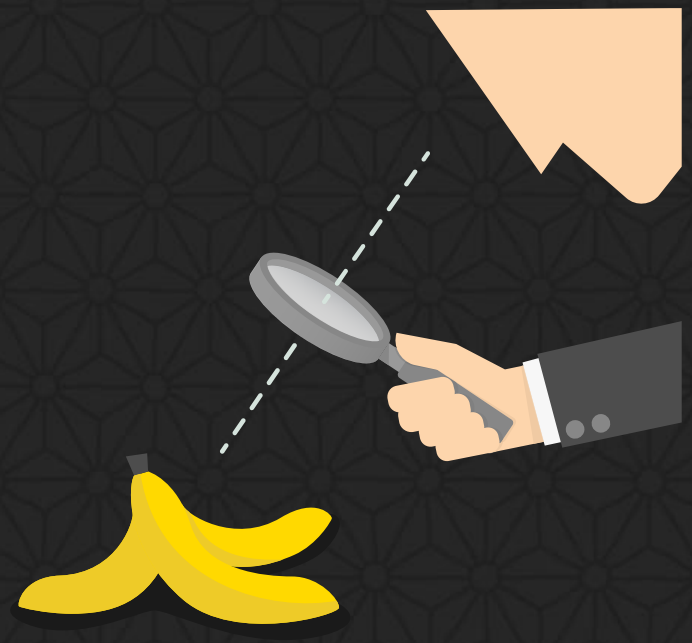
[1] https://www.thinkwithgoogle.com/consumer-insights/voice-assistance-consumer-experience/

[2] Laughter’s Influence on the Intimacy of Self-Disclosure, Gray, A.W., Parkinson, B. & Dunbar, R.I. Hum Nat (2015) 26: 28.

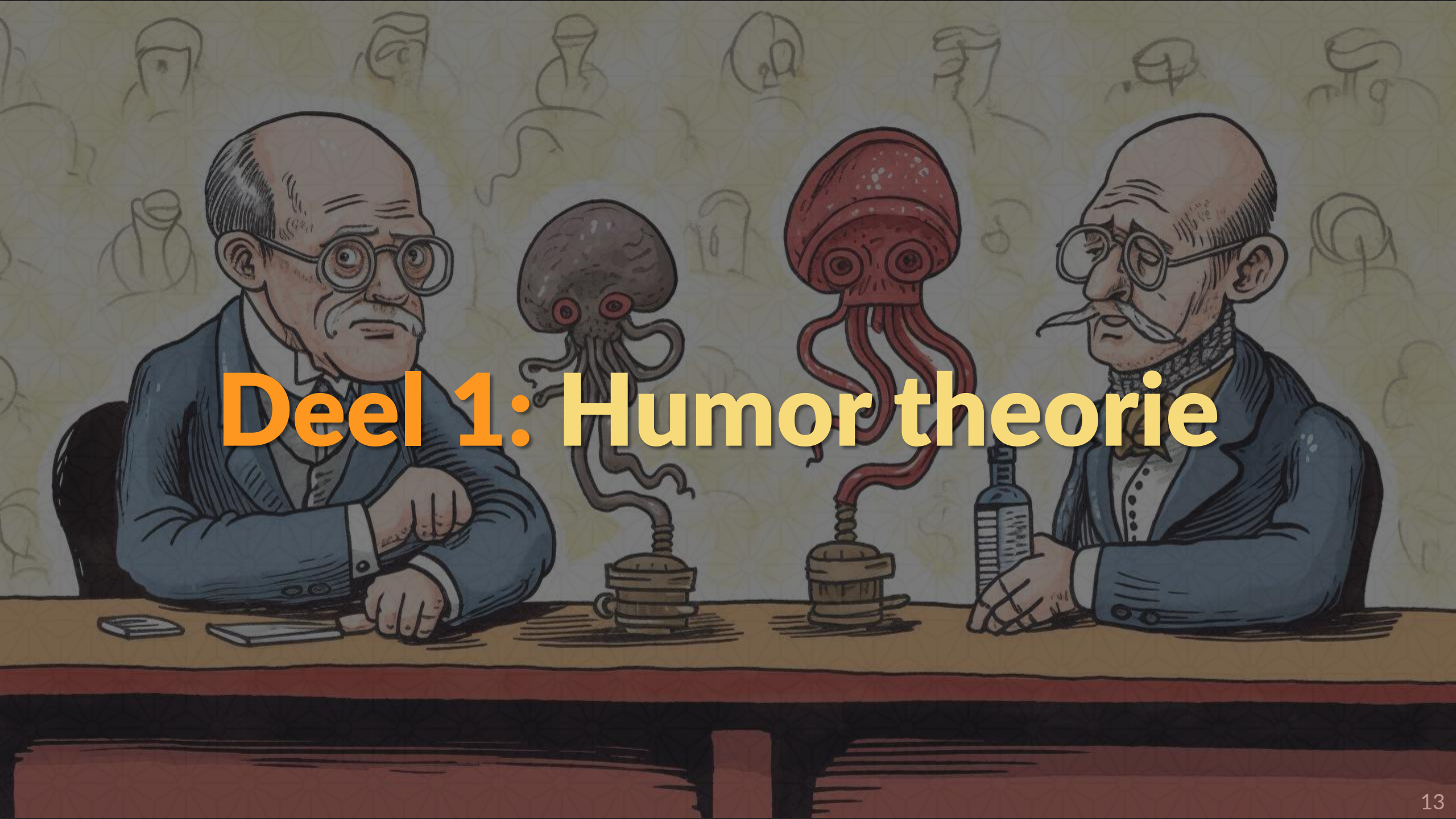
Reden 3: Bestuderen van humor theorie

We weten nog altijd niet
helemaal **hoe humor werkt**





Artificiële Intelligentie kan helpen
meten & begrijpen!



Deel 1: Humor theorie

Waarom hebben mensen humor?

Brein beloont opmerken & oplossen van **incongruenties**

\+ teken van **intellectuele** & linguistische vaardigheden

\+ moeilijk vervalsbaar vertoon van **persoonlijke waarden**

= **Evolutionair voordeel!**



Incongruëntie-Resolutie Theorie

Setup *Ik wil sterven zoals mijn opa: rustig in zijn slaap.*

Punchline *Niet zoals zijn schreeuwende passagiers*

Obvious Interpretatie



Verborgen Interpretatie



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

Mens-gefocuste definitie

Computer moet niet alleen de **twee mentale beelden opmerken**

Obvious Interpretatie



Verborgen Interpretatie



Maar ook dat ze voor mensen niet **te moeilijk of te makkelijk zijn**!

Humor ≡ AI-compleet?

~ even moeilijk als computers zo “intelligent” als mensen maken

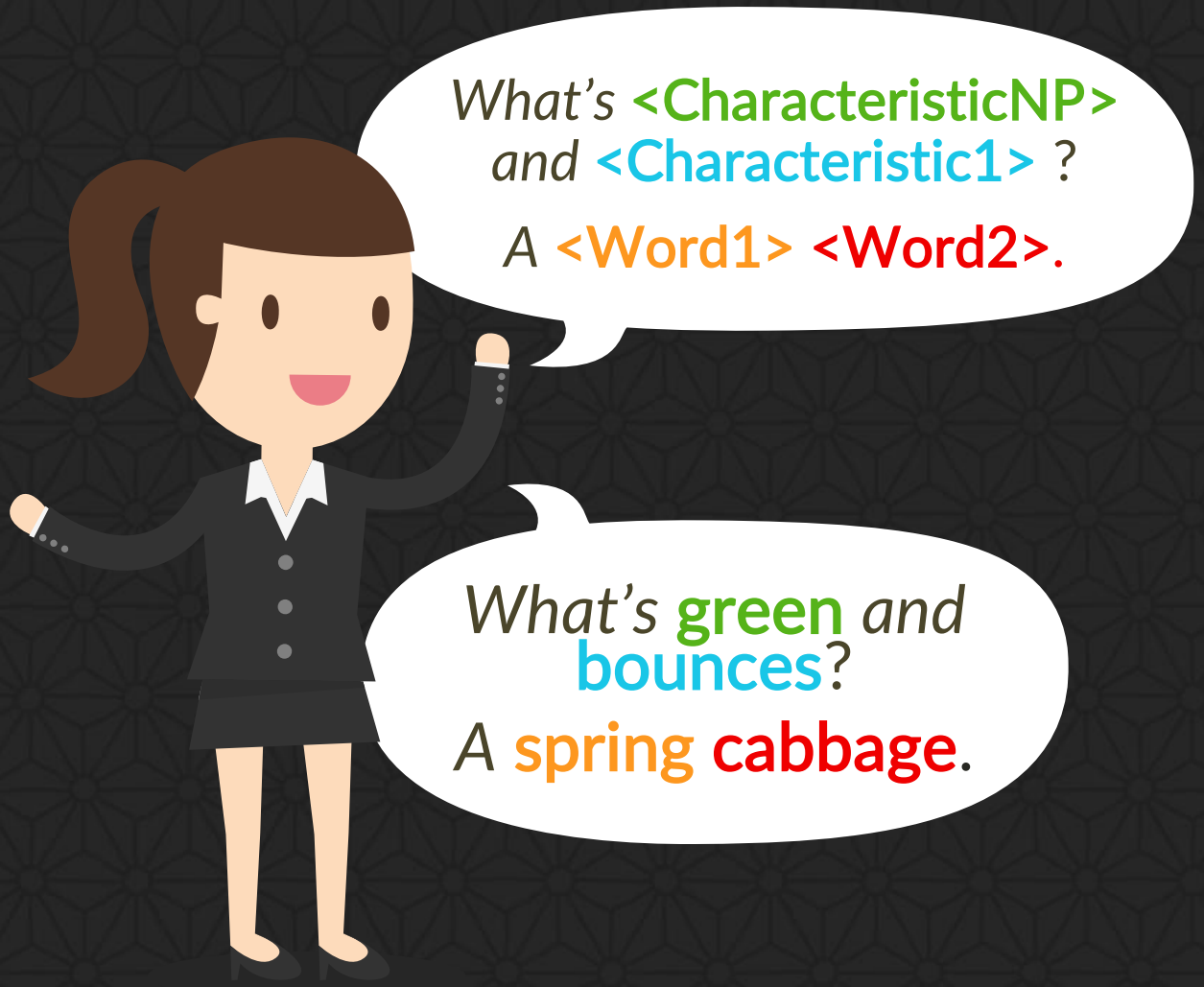


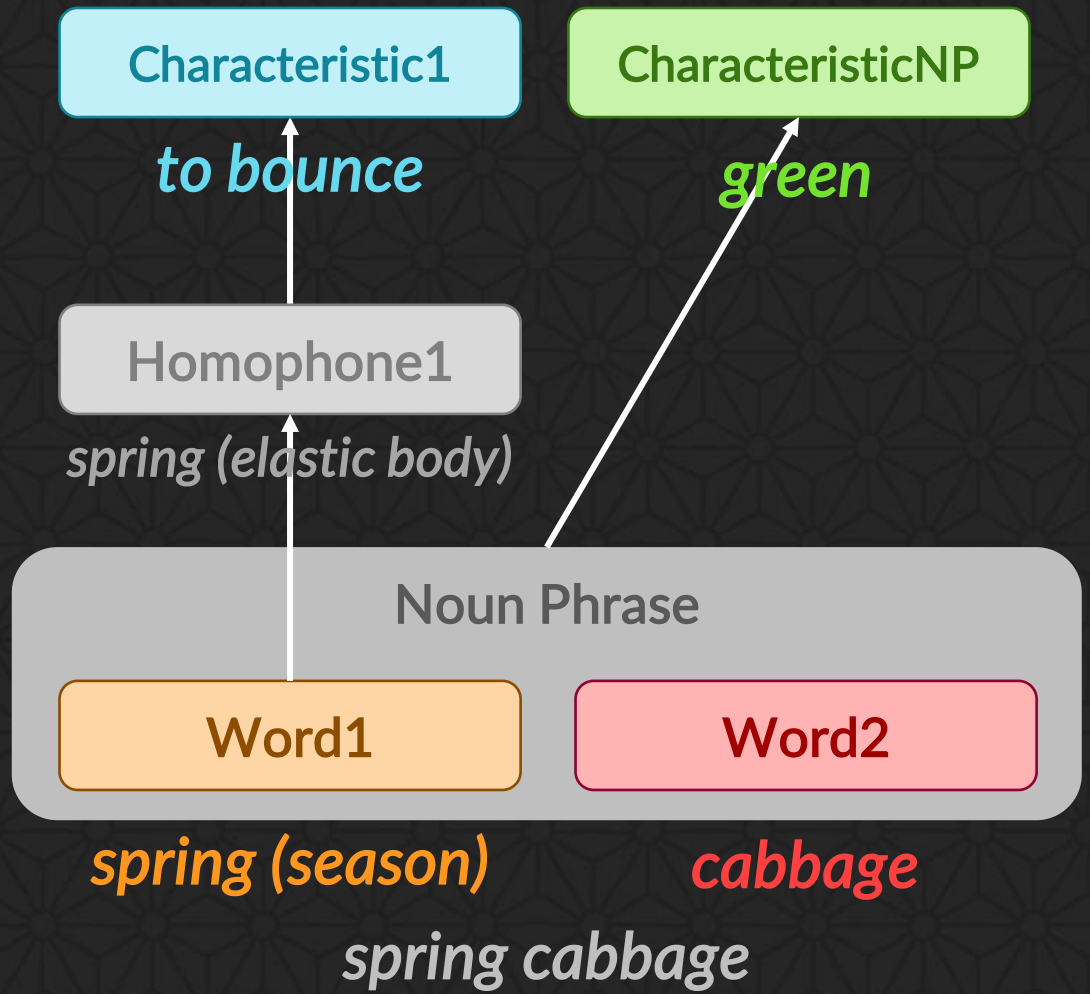
Deel 2: Hoe schrijven computers humor?



Manier 1:
Schrijf zelf wat mopjesregels

Eerste humor generator: Sjablonen & Schema's





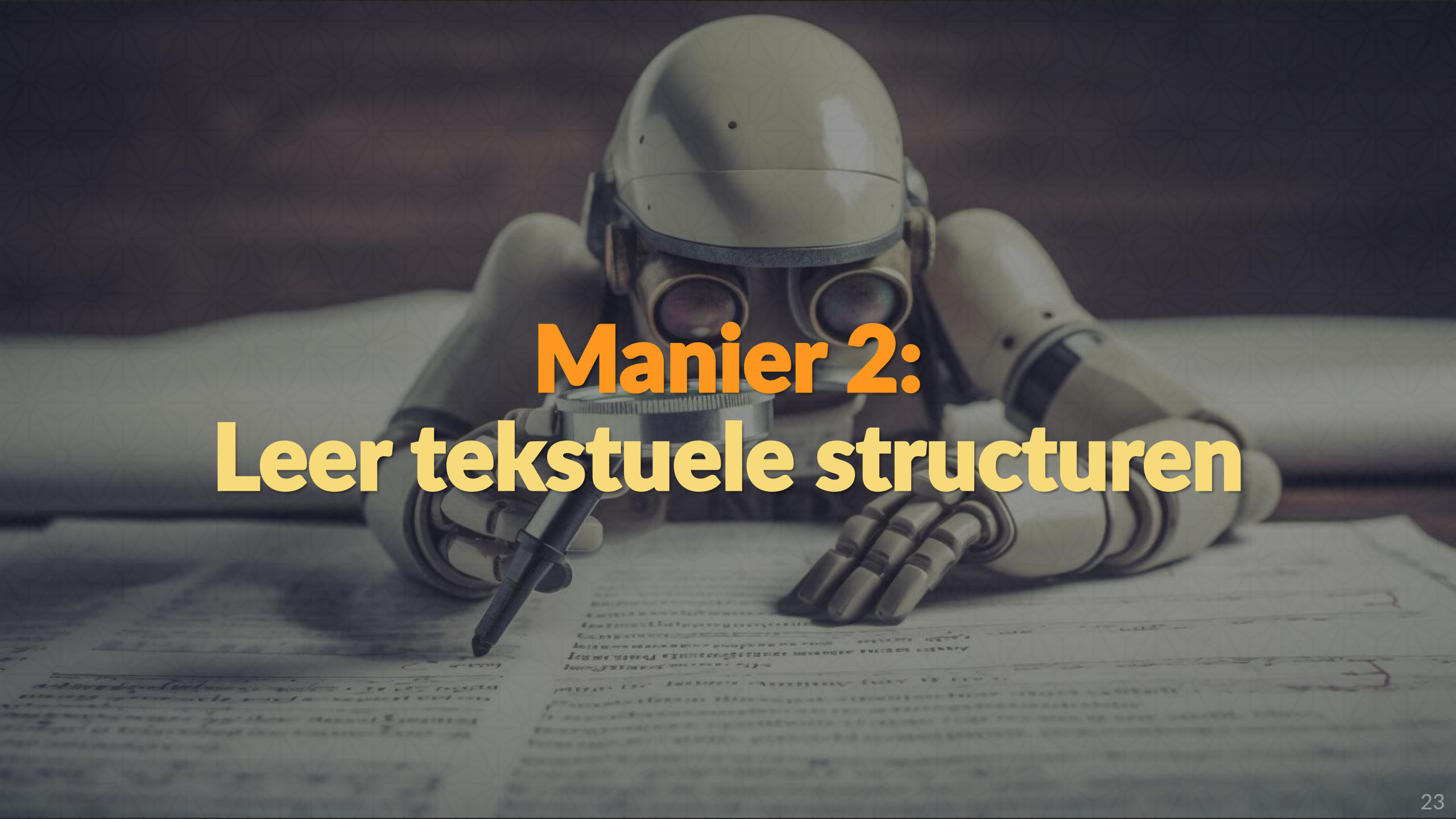


https://www.youtube.com/watch?v=X0dl_y41zdo

MopjesBot


Het is een <Beschrijving1>
en <Beschrijving2> ?
<EersteDeel> <Rijmwoord>.
Het is een Belgische politica
en komt tot net boven de
enkel?
Maggie De Sok.


Beschrijving2
Komt tot net
boven de enkel
Beschrijving1
Rijmwoord
Belgische
politica
Sok
Naam
EersteDeel
LaatsteDeel
Maggie De
Block
Maggie De Block



Manier 2:
Leer tekstuele structuren

Tel hoe vaak woorden na andere komen

1. Tel in alle Rik Torfs teksten hoe vaak een woord volgt op de vorige paar woorden



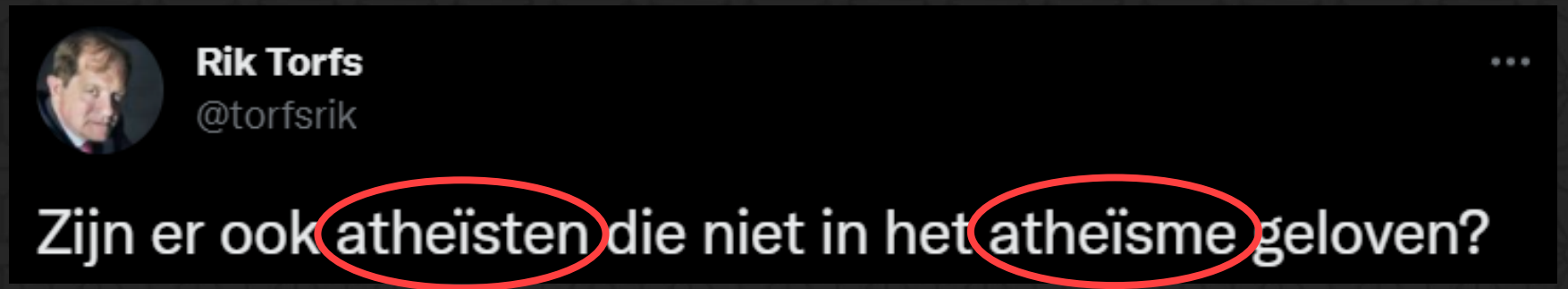
2. Begin met beginwoorden van Rik & neem steeds willekeurig mogelijk volgend woord





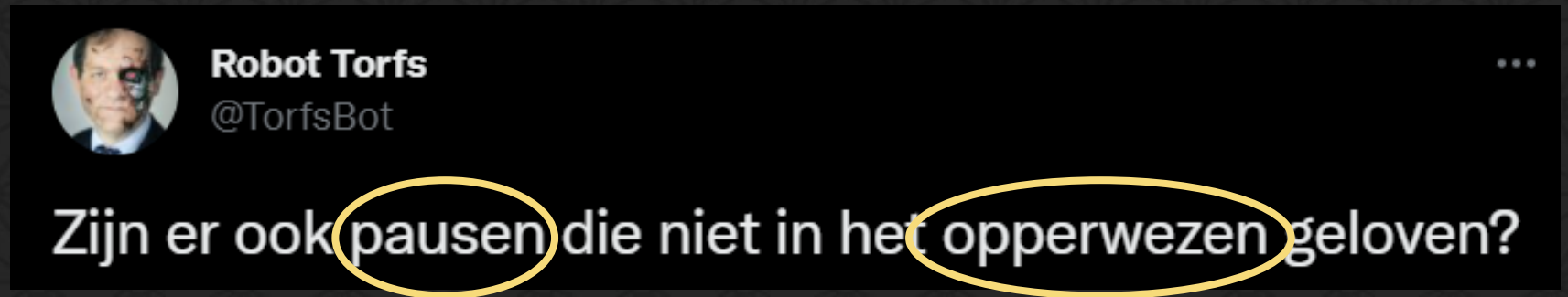
OF: neem patroon over door kernwoorden te vervangen

Neem tweet & column



Rik Torfs Column: *(...) Het feit dat het voormalige Opperwezen dit nieuw verworven inzicht niet probeert te ontkennen, zien zij als een bewijs van hun gelijk. Ook met de Kerk zit het niet snor. Norse pausen.*

Vervang kernwoorden van tweet door zelfde type woorden




Manier 3:
Leer humor herkennen

Eerste humor detectors (2005)

Gebruikte simpel AI model + simpele humorkenmerken

(alliteratie, volwassen jargon...)

Taak: onderscheid one-line mopjes van nieuws, neutrale tekst en gezegden

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.

Maar is dit wel écht humordetectie?

Nieuws & gezegden hebben compleet **verschillende soorten woorden & lengte** dan mopjes!





→ AI kan valsspelen met woordfrequenties in datasets

Mopjes zijn **fragiel**!

"Ober, kunt u die schrik uit mijn politie man halen?
Want ik eet liever alleen."

Maak kapotte mopjes door
kernwoorden te veranderen (~ TorfsBot)

Woordfrequentie trucjes als onderscheid werken nu niet meer!

Voorbeelden van kapotte mopjes

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.

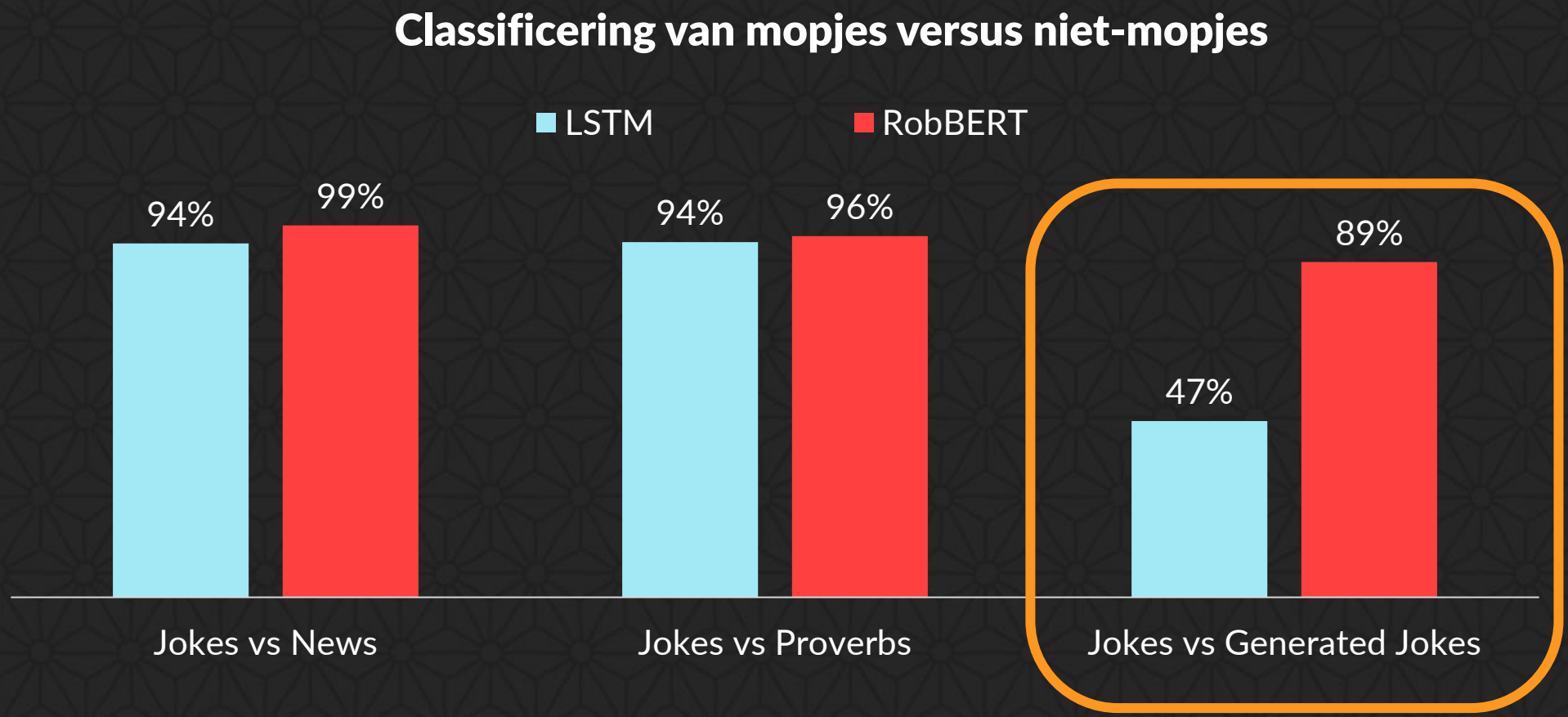
Er loopt een super vriendelijk blondje langs een armband. Last er een toonbank: “zo, waargaan die mooie mannen heen?” Blondje: “naar de barkeeper als er niets tussen komt…”

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!'"



Veel moeilijkere dataset



Échtere humordetectie?



Manier 4:
Vraag het lief aan GPT

GPT-4 mopjes

USER	Schrijf een korte mop over het Nerdland Festival
ASSISTANT	Waarom kon het Nerdland Festival niet doorgaan dit jaar? Omdat alle deelnemers te druk waren met het updaten van hun software!

Probleem

GPT voorspelt “logische” volgende woorden

VS

Mop breekt verwachting



Net zoals GPT “slecht is in wiskunde”?

Thomas takes a train from Leuven to Gent. Leuven Station and Gent have a distance of 100km. The train travels at 50km/h. The train also stops three times for 10 minutes each. How long is the train ride?

2 hours

“Let’s think step by step”

Thomas takes a train from Leuven to Gent. Leuven Station and Gent have a distance of 100km. The train travels at 50km/h. The train also stops three times for 10 minutes each. How long is the train ride? Let's think step by step.

The train travels at 50km/h, so it will take 2 hours to travel 100km. However, the train stops three times for 10 minutes each. This means that the train will spend a total of 30 minutes stopped. Therefore, the total train ride will be 2 hours and 30 minutes.

USER

You are a world-renowned expert in comedy with a proven track-record in writing jokes and stand-up comedy for nerdy audiences. You task is to write a list of jokes related to the Nerdland Festival, an open-air science communication festival with talks and workshops instead of musical acts. You write every joke by first thinking step by step.

ASSISTANT

Joke 1:
Step 1: Identify a nerdy topic - Quantum physics
Step 2: Relate the topic to the festival setting - Workshops and talks
Step 3: Set up the punchline - Misunderstanding of quantum concepts

Joke: Why are the quantum physics workshops the most popular at Nerdland Festival? Because they're both in session and not in session at the same time! Schrödinger would be proud.

Joke 2:
Step 1: Identify a nerdy topic - Computer programming
Step 2: Relate the topic to the festival setting - Networking with other attendees
Step 3: Set up the punchline - Social interaction difficulties

Joke: How can you tell the difference between a programmer and a social butterfly at Nerdland Festival? The programmer tried to start a conversation with a "Hello, World!" t-shirt.

Je kan zelf stappen specifiëren

Je bent een wereld-expert in comedy met een bewezen track record in het schrijven van hilarische one-liners in het Nederlands. Ik zoek naar goede mopjes voor gegeven topics. Je taak is om stap voor stap een grappige mop te schrijven door brainstorms van associaties en redeneerprocessen neer te schrijven, om dan enkele punchlines te bedenken voor een topic.

Altijd als je een mop schrijft over een onderwerp, volg je de volgende stappen:

1. Onderwerpen: lijst de sleutelwoorden op als kommalijst

2. Associaties: Brainstorm een lijst aan associaties aan de onderwerpen uit stap 1 als komma lijst.

3. Link: Schrijf een grappige link tussen twee associaties uit stap 2.

4. Punchlines: Schrijf drie punchlines op basis van de link gevonden in stap 3.

5. Beste punchline: selecteer de beste punchline uit stap 4.

6. Schrijf drie setups die passen bij de punchline uit stap 5

7. Volledige mop: selecteer de beste setup uit stap 6, en maak een vloeiende overgang naar de beste punchline uit stap 5.

USER

Schrijf een mop over Nerdland Festival, een festival met wetenschapstalks en workshops in plaats van muziekoptredens.

ASSISTANT

1. Onderwerpen: Nerdland Festival, wetenschapstalks, workshops, muziekoptredens

2. Associaties: nerds, professoren, experimenten, lezingen, dansvloer, festivalterrein, laboratorium, tenten, drank, eten, Einstein, hoofdtelefoon

3. Link: De nerds hebben een festival waar ze kunnen drinken en hun innerlijke Einstein loslaten in plaats van losgaan op de dansvloer.

4. Punchlines:
A) Nerdland Festival: het feest waar science op proef wordt gesteld en het bier wordt geanalyseerd.
B) Op het Nerdland Festival treffen we de meest flitsende natuurkundigen die zowel het Higgs-deeltje als de akoestische eigenschappen van een Red Bull-blikje kunnen uitleggen.
C) Wil je op het Nerdland Festival alcohol krijgen? Dan moet je je bestelling afwegen in mol!

5. Beste punchline: Wil je op het Nerdland Festival alcohol krijgen? Dan moet je je bestelling afwegen in mol!

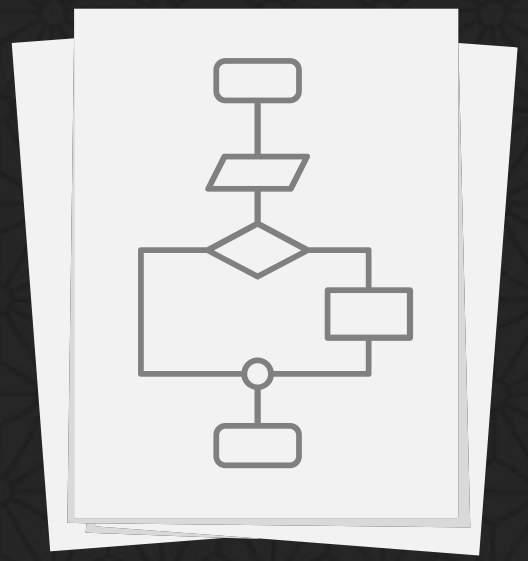
6. Setups:
A) Weet je wat het verschil is tussen een gewoon festival en het Nerdland Festival?
B) Heb je ooit gehoord van het Nerdland Festival? Het is niet jouw gewone feestje!
C) Je zou niet geloven wat er gebeurt op het Nerdland Festival als je een biertje wilt bestellen!

7. Volledige mop: Weet je wat het verschil is tussen een gewoon festival en het Nerdland Festival? Wil je op het Nerdland Festival alcohol krijgen? Dan moet je je bestelling afwegen in mol!

Kunnen computers een gevoel voor humor hebben?



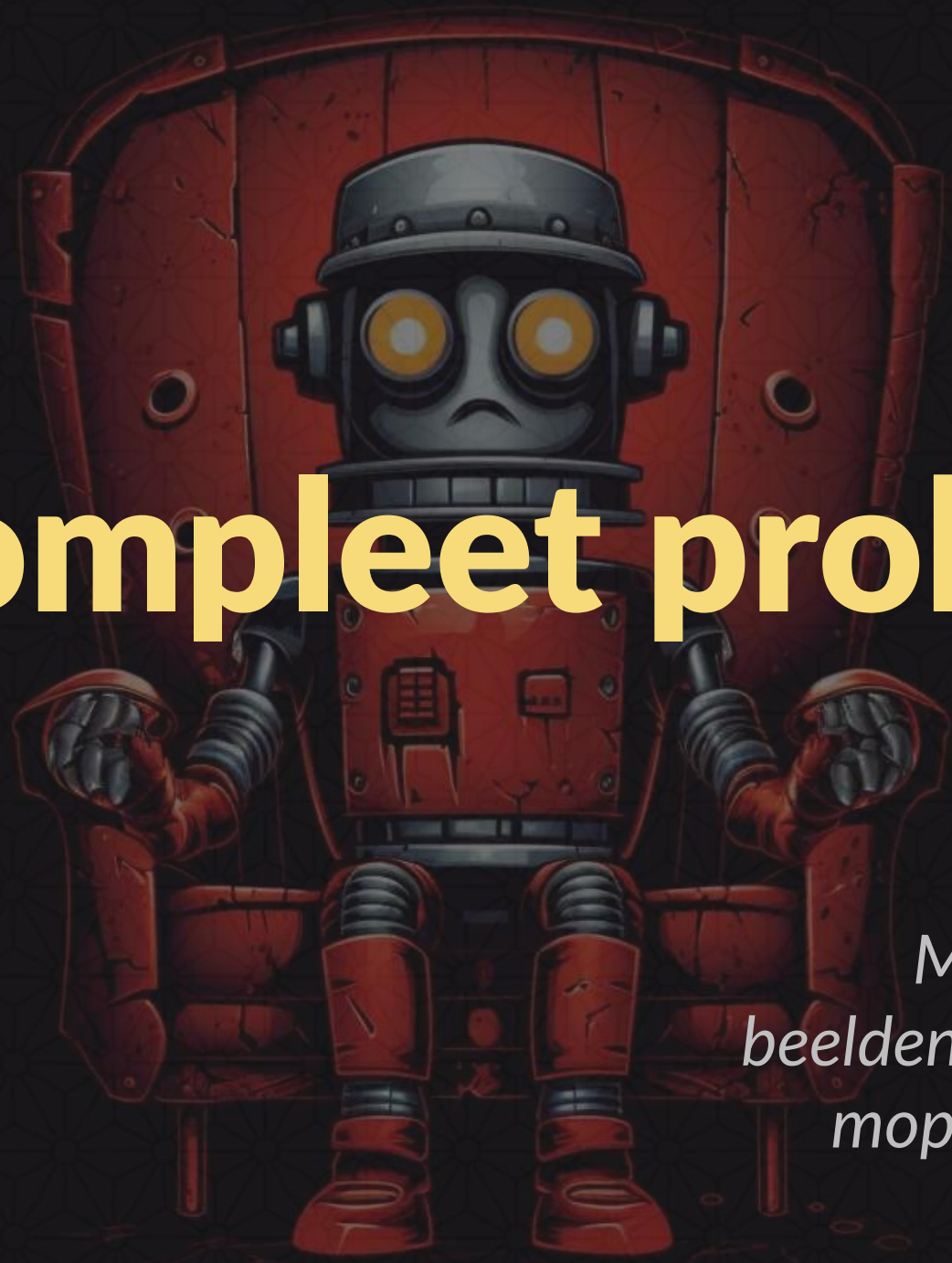
Humor is **intrinsiek menselijke**



Computers kunnen **specifieke humor instructies volgen**

Grote taalmodellen leiden de revolutie

maar:



AI-compleet problem!

Moet de twee mentale beelden snappen om écht de mop te kunnen begrijpen.

Hoe schrijven computers humor?



Thomas Winters

PhD Student at KU Leuven & FWO Fellow

@thomas_wint

thomaswinters.be

Slides: thomaswinters.be/talk/2023nerdlandfestival