GENERATING DUTCH PUNNING RIDDLES ABOUT CURRENT AFFAIRS

Thomas Winters
KU Leuven
@thomas_wint

VIRTUAL ASSISTANTS

People often treat virtual assistants like humans

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

41%

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

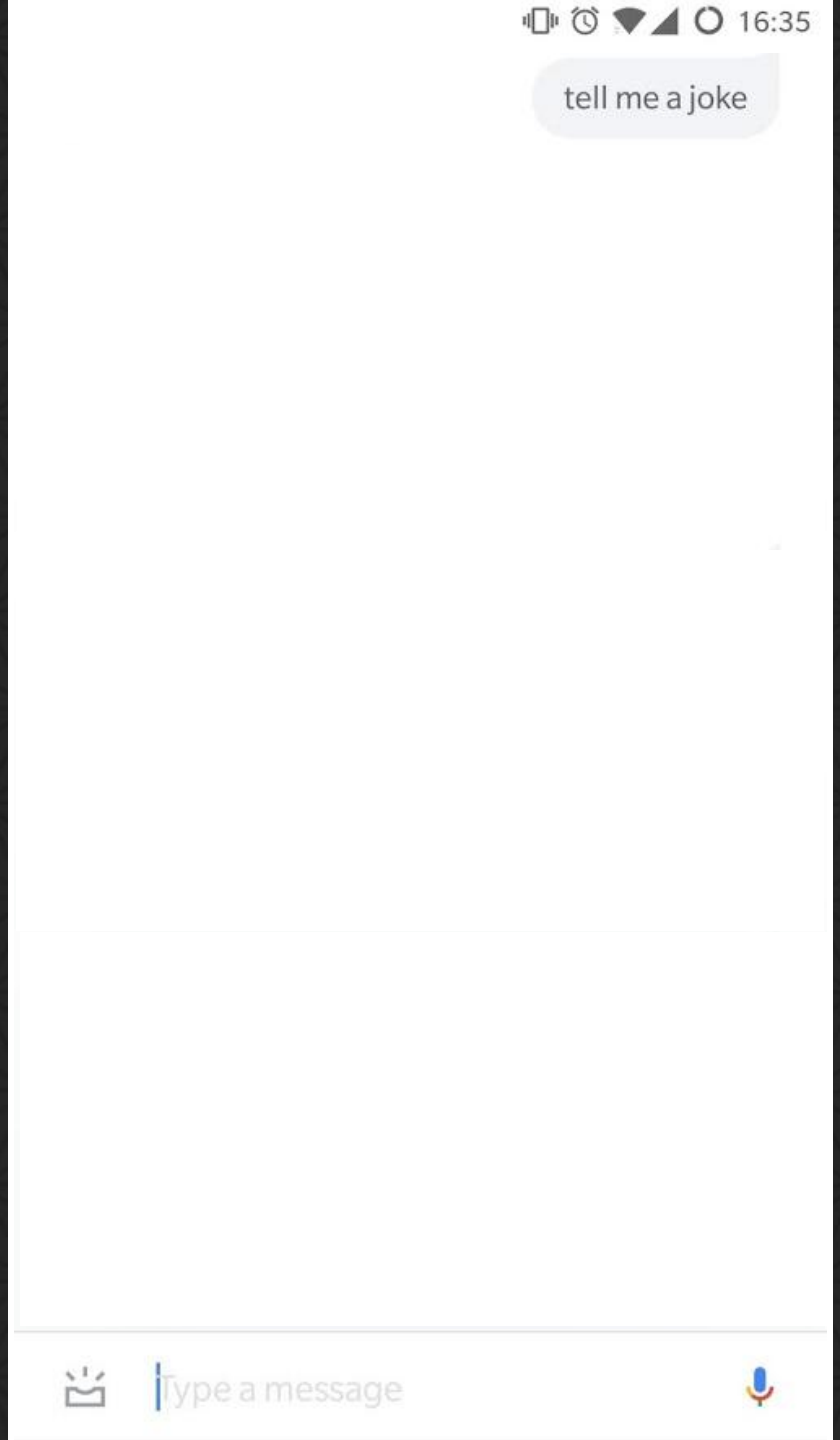
Humor and creativity in language is important **between friends** [2]



[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.

CURRENT ASSISTANT HUMOUR



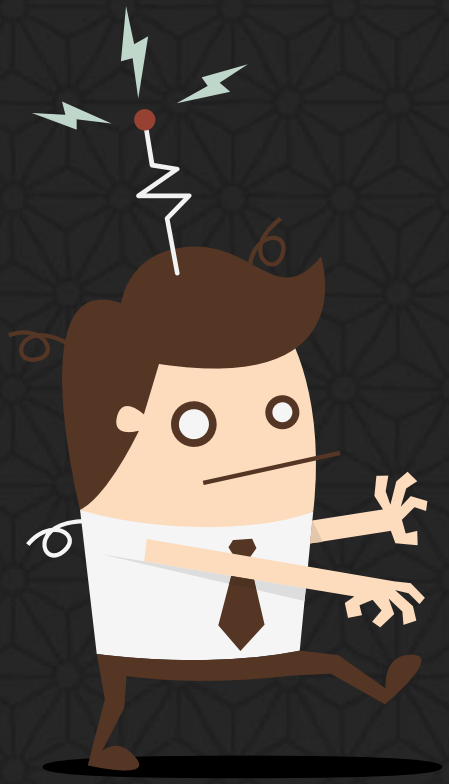
NICE, BUT:



PREWRITTEN JOKES

→ Limited # jokes

→ No input: actual events, user (friends)…



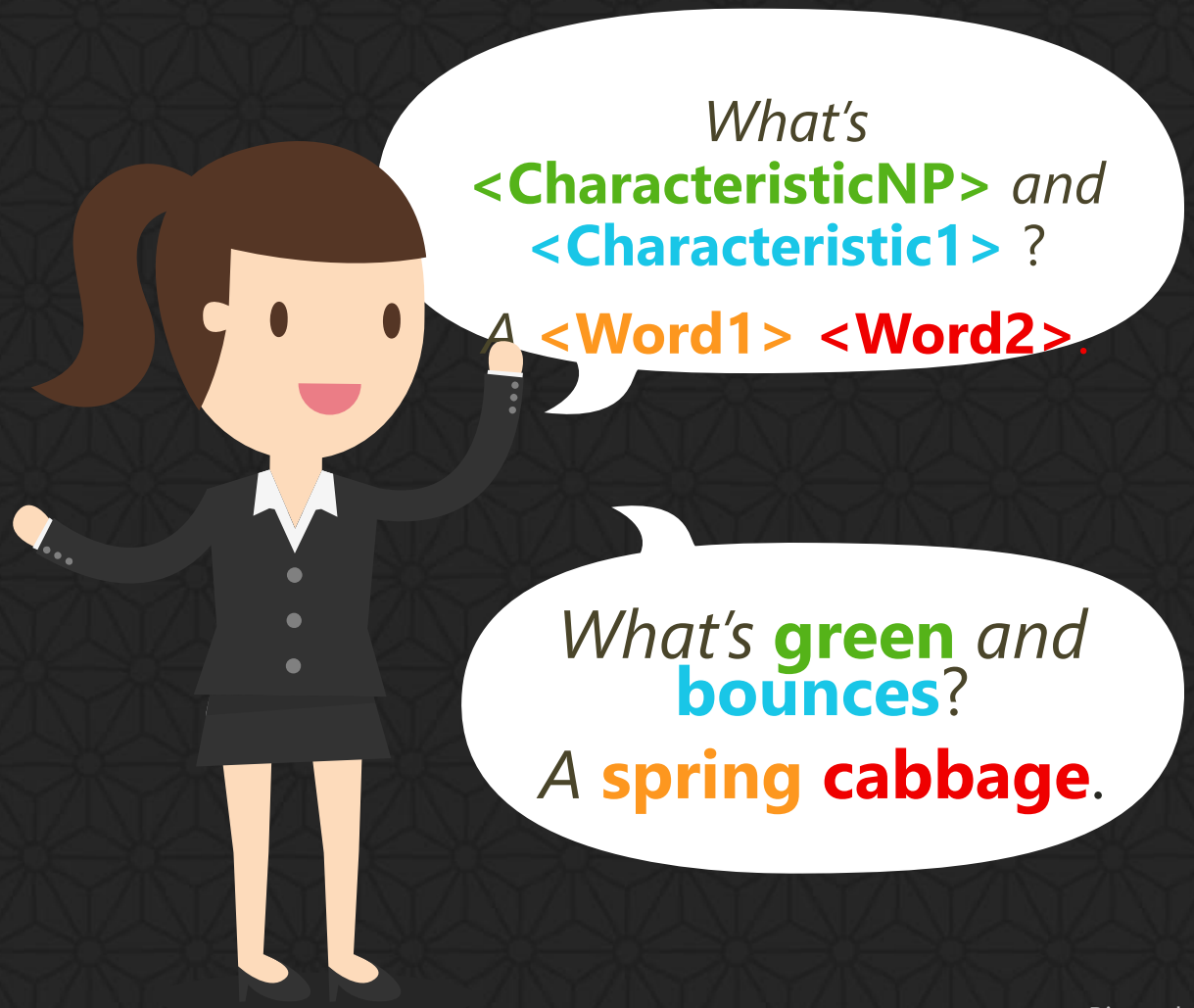
AUTOMATION?

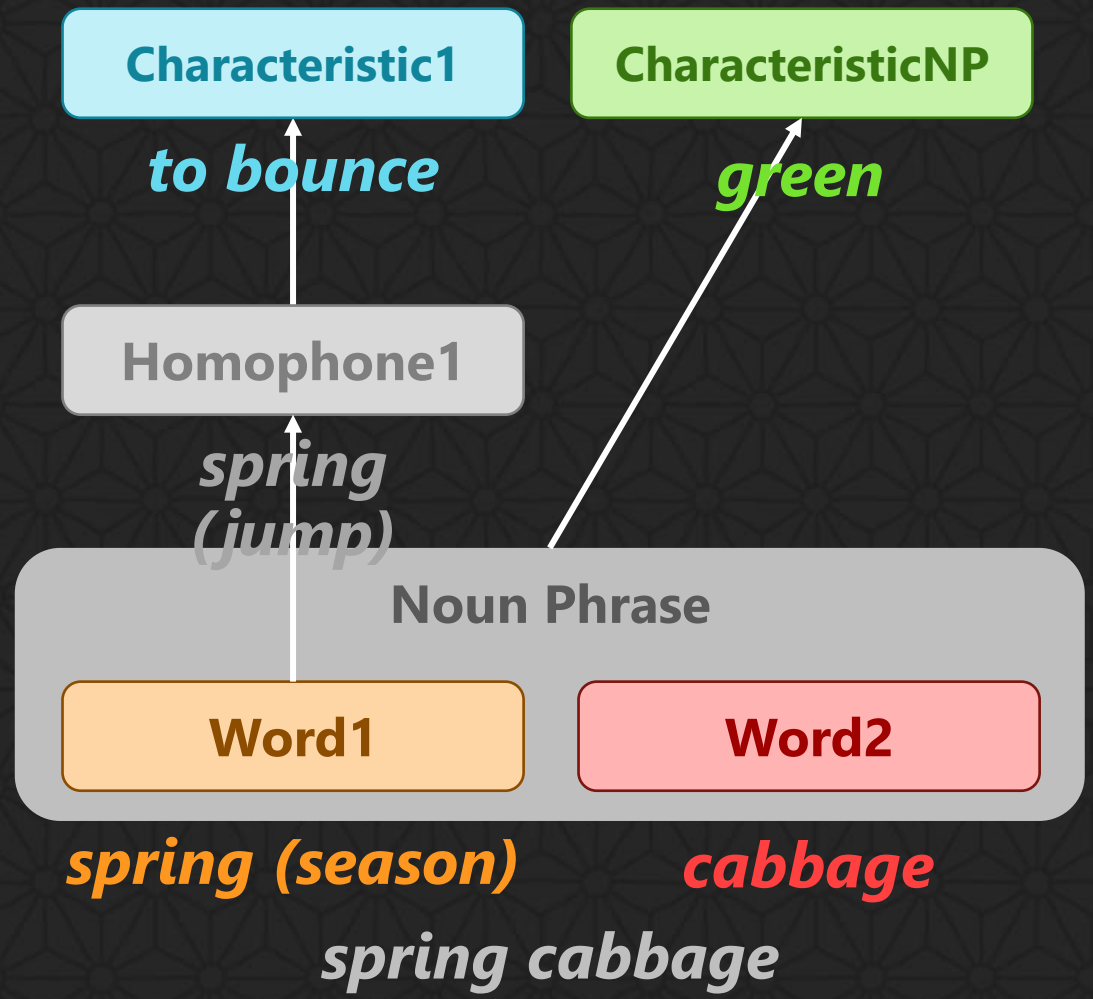
Riddle pattern generatable:

about any topic?

EXISTING WORK: JAPE & STANDUP

Uses templates and schemas with knowledge base





Binsted, K., & Ritchie, G. (1994). An implemented model of punning riddles. In proceedings of AAAI-94.

Ritchie, G., Manurung, R., Pain, H., Waller, A., & O'Mara, D. (2006). The standup interactive riddle-builder. IEEE Intelligent Systems, 21(2), 67-69.

Manurung, R., Ritchie, G., Pain, H., Waller, A., O'Mara, D., & Black, R. (2008). The construction of a pun generator for language skills development. Applied Artificial Intelligence, 22(9), 841-869.

MISSING COMPONENTS JAPE & STANDUP



WORLD TRIVIA

(WordNet-like)



CURRENT AFFAIRS

(static knowledge base)



PERSONALISATION

(no jokes about yourself, friends or input words)



DUTCH

(only English)

SORRY VOOR ALLES: ATTENTION TO JOKE PATTERN



Youtube: https://www.youtube.com/watch?v=X0dl_y41zdo
Source: https://www.een.be/sorry-voor-alles/daten-met-een-moppentapper (4m30->5m10)

LET'S ANALYSE

Het is groen en het plakt aan de muur?

Kermit de Sticker

Kermit de Sticker
Kermit de Pikker
Kermit de Likker
~~*Kermit de Keeper*~~
~~*Kermit de KikKerk*~~
~~*Kermit de Verklikker*~~

INTERESTING NEW CONSTRAINTS

RHYME BASED
SAME # SYLLABLES

JAPE/STANDUP VS OUR WORK



WORLD TRIVIA & OTHER LANGUAGES

↓

Wikipedia, Wiktionary *(trivia)*
& mijnwoordenboek.org



CURRENT AFFAIRS

↓

Online news sites
(VRT NWS, Nieuwsblad…)



PERSONALISATIO N

↓

Social media
(Twitter bio)

SIMPLE PUN FOR NAME

1. Get name from input
2. Find last word
3. Normalize Dutch last name spelling
4. Split into syllables and count *(hyphenate)*
5. Find rhymes *(rhyming dictionary)*
6. Filter out rhymes with more syllables
7. Choose common word *(unigrams)*
8. Insert rhyme, respecting syllables

1. Kermit De Kikker
2. Kikker
3. Kikker
4. Kik-ker (2)
5. bikker, ketelbikker, knikker, steenbikker, dikker, pikker…
6. bikker, knikker, dikker, pikker…
7. Pikker
8. Kermit De Pikker

RIDDLE: DESCRIBING NAMED ENTITIES

1. Search on Wikipedia, take first paragraph
2. Search for first word with stem "to be"
3. Take words until punctuation, start of clause or propositions

Kermit de Kikker

Kermit de Kikker is een muppet die in de jaren vijftig werd bedacht door Jim Henson en nadien een belangrijke rol kreeg in diverse series en films. Hij is een groene muppet die een kikker voorstelt.

Martin Luther King

Martin Luther King, oorspronkelijk *Michael Luther King, Jr.* (Atlanta (Georgia), 15 januari 1929 – Memphis (Tennessee), 4 april 1968), was een Amerikaanse baptistendominee, politiek leider en een van de prominentste leden van de Afro-Amerikaanse burgerrechtenbeweging.

Verenigde Staten

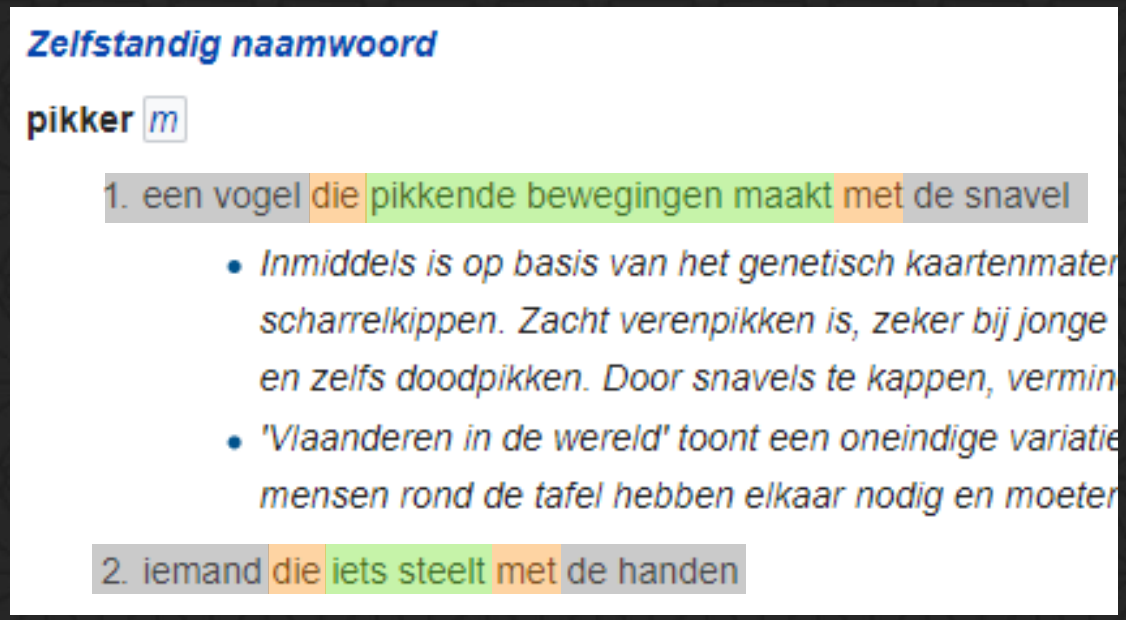
Beluister

Zie Verenigde Staten (doorverwijspagina) voor andere betekenisse

De **Verenigde Staten**, officieel de **Verenigde Staten van Amerika**, afgekort **VS** (Engels: *United States of America*, afgekort *USA*, *US of A* of *US*), vaak (totum pro parte) **Amerika** (*America*) genoemd, zijn een federatie van 50 staten en het District of Columbia, grotendeels in Noord-Amerika gelegen. De Verenigde Staten zijn met meer dan 325 miljoen inwoners het op twee na grootste land ter wereld in bevolking, na Volksrepubliek China en India. Tevens is het het op twee na grootste land in oppervlakte, na Rusland en Canada.[noot 1]

RIDDLE: DESCRIBING PUN WORD





JAPE&STANDUP: SYNONYMS

We can use Dutch synonym database

ACTIONS GENERALLY FUNNIER

Extract action from definition

GENERATE THE RIDDLE

1. Generate pun on name
2. Use Wikipedia to describe first part
3. Use (puzzle) synonyms, Wiktionary or mijnwoordenboek.org for second part
4. Pick non-spoiler, common description (least frequent of unigrams)
5. Put descriptions into template

1. Kermit De Kikker + pikker = Kermit De Pikker
2. *Kermit de Kikker is* **een muppet** die *in de jaren vijftig*
3. *iemand die* **iets steelt** *met de handen* OR *dier, gauwdief, rover, steler, vogel…*
4. *verb: "stelen", object: "iets".*
5. Het is een muppet en het steelt iets?

IMPROVEMENTS ON BASIC ALGORITHM



Het is een muppet en het is hem bij het spel?

Kermit de Tikker!

Het is een muppet en het is een korte niet al te harde klap?

Kermit de Tikker!

Het is een muppet en is een oogopslag?

Kermit de Blikker!

JOKE SERIES

Makes it feel like a puzzle & clears up name

DIFFERENT SYLLABLES

Clears up target

ORIGINALITY

Sometimes same pun found in different way: avoid!

MAKING IT ABOUT CURRENT AFFAIRS



For all news sites:

1. Scrape all articles
2. Filter articles with sensitive topics
3. Extract all named entities (capital letters)
4. Bundle together name counts of same person

1. Rank names from high to low # occurrences
2. Filter out names used in recent past
3. *Optional: filter with Wikipedia page (= public figure)*
4. Take name with highest occurences

FULL ALGORITHM





Rhyming
Hyphenation
Unigrams

Used Puns
Joke Target
Database

Wiktionary
Wikipedia
Mijnwoordenboek.org

PUN GENERATOR

ORIGINALITY CHECKER

RIDDLE GENERATOR

Het is een Nederlandse podiumkunstenaar en is helemaal niets?
Herman van Geen

Herman van Veen

Herman van Geen
Herman van Steen

Herman van Geen

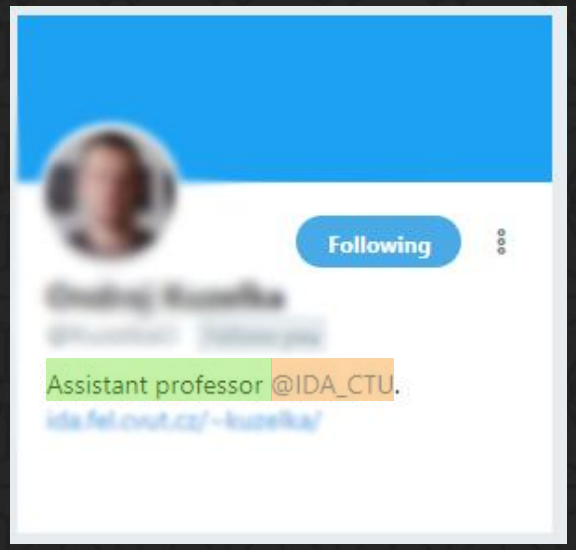
- Nederlandse Podiumkunstenaar
- Helemaal niets

JOKES ABOUT NON-PUBLIC FIGURES

Twitter users usually describe themselves well already in their **bio**:

Short description usually ends with **punctuation** or **tag**





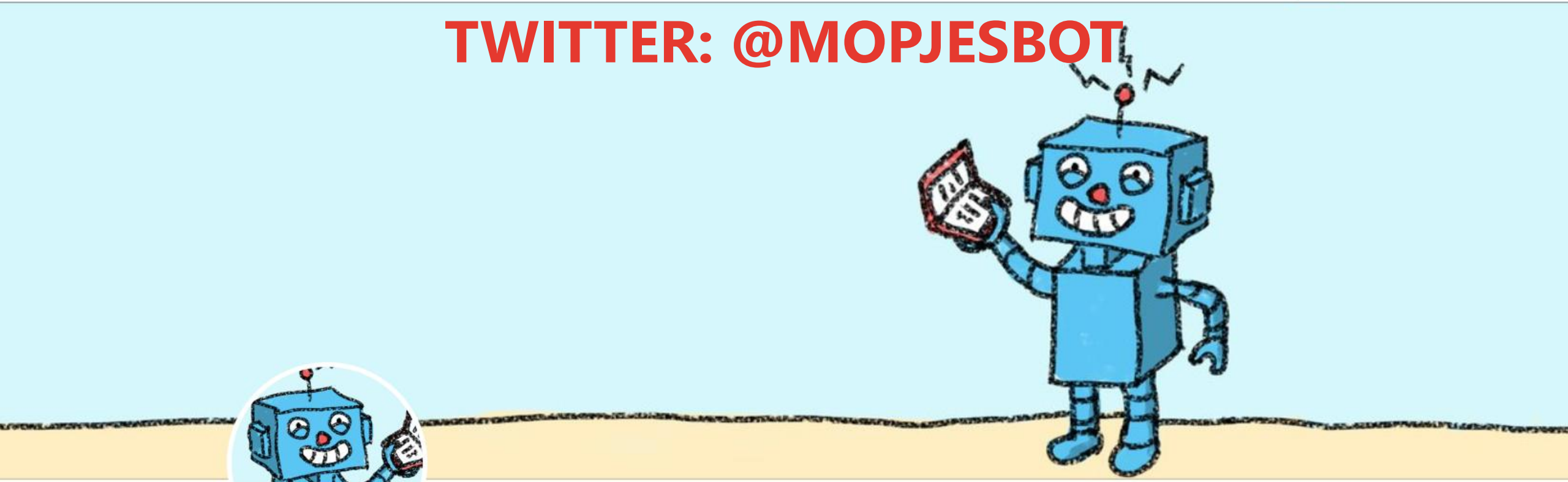




Otherwise, just use **first name**: *"Het is een Thomas en ..."*

Home Notifications Messages Search Twitter Tweet

TWITTER: @MOPJESBOT



Tweets 337 Following 1 Followers 25

Following

MopjesBot

@MopjesBot Follows you

Ik ben een robotje dat elke dag zijn uiterste best doet om woordspelingen over het nieuws te maken! (door @thomas_wint)

Joined September 2018

Tweet to Message

Tweets Tweets & replies

MopjesBot @MopjesBot · 3h
Het is professor en is een nederlandse politieke figuur?

Paul De Pauwe!

Translate Tweet

1

Show this thread

MopjesBot @MopjesBot · Jan 26
Het is een Amerikaans regisseur en producent en is sar?

Steven Spielterg!

SOME EXAMPLES

Het is een kinderprogramma en is een grap?
Kabouter Mop!

Het is een kinderprogramma en bedoelt om iemand mee op te hangen?
Kabouter Strop!

Het is een kinderprogramma en is geweldig?
Kabouter Top!



FUTURE /CURRENT WORK

Het is een muppet en het is een succesvol lied?
Ker-hit de Kikker

Het is een muppet en het is afgelegen?
Ver-kit de Kikker

PUN GENERATOR 2
"Banaanhangwagen"



DIFFERENT SYLLABLES

Trivial expansion, but
- breaks "puzzle pattern" if not careful
+ makes pun target more clear

SWAP PUN GENERATOR

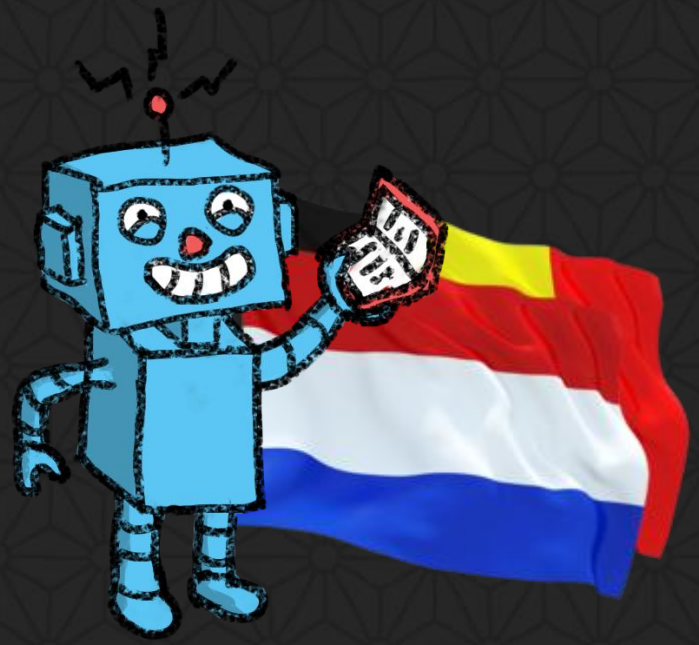
E.g. for "Banaanhangwagen" puns. *(WIP)*

DBPEDIA DESCRIPTORS

Can create more *"uniquely identifiable"* descriptors, although original joke is vague itself (groen). Similar to RiddlerBot *[4]*

[4] Román, Iván Guerrero et al. "TheRiddlerBot: A next step on the ladder towards creative Twitter bots." *ICCC* (2015).

CONCLUSION



DUTCH PUNNING RIDDLES

Possible to generate fun & funny riddles in Dutch with current tools & data bases



CURRENT AFFAIRS

Joke generation about people and other trending news entities



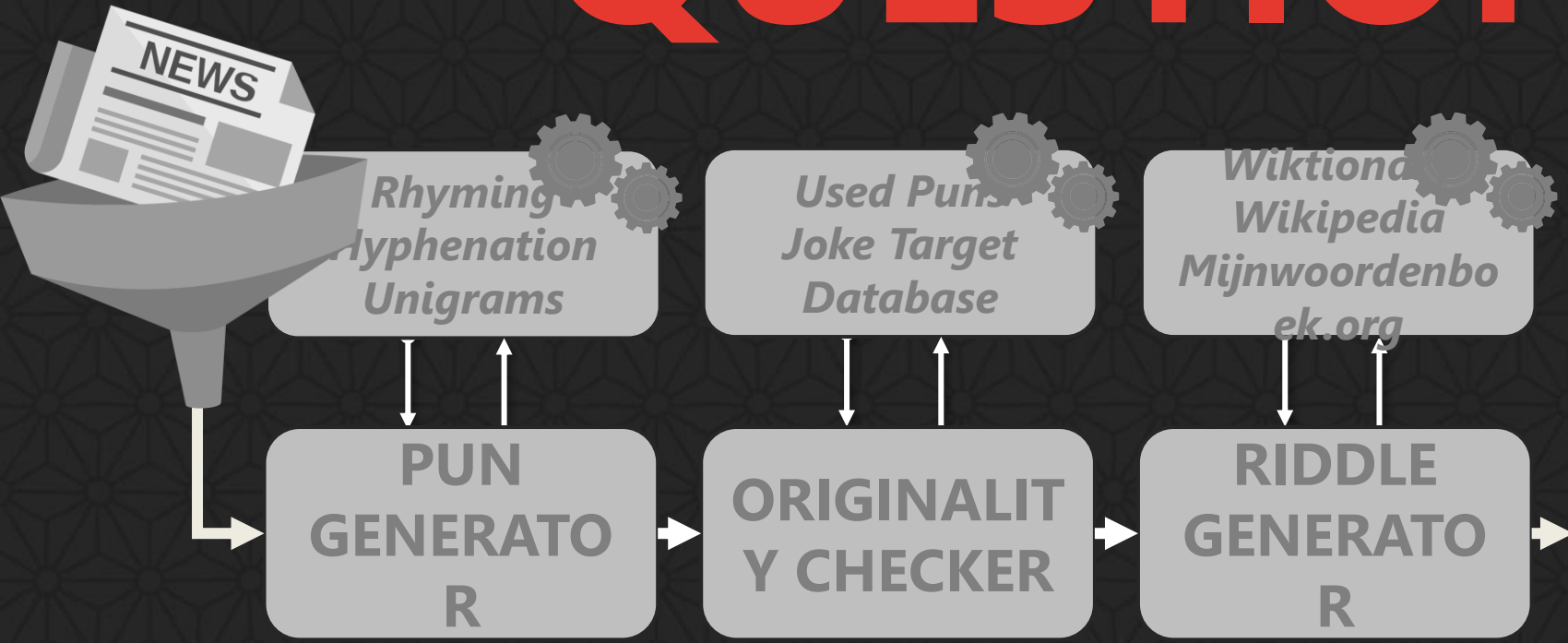
SOCIAL

Allows joke generation about user & friends: ~interactive

THANK YOU!
QUESTIONS?

Thomas Winters
KU Leuven
@thomas_wint

twitter.com/MopjesBot
@MopjesBot



Het is een Thomas en het drukt gegevens af onder andere op papier?
Thomas Printers!

Het is een Thomas en is soort jenever?
Thomas Ginters!

Het is een Thomas en kan heel snel korte afstanden afleggen?
Thomas Sprinters!

CREDITS & REFERENCES

Video

- VRT: Sorry voor alles

Images

- pexels.com
- Business vector created by dooder - www.freepik.com
- www.flaticon.com
- Business vector created by macrovector - www.freepik.com
- Arrow vector created by rawpixel.com - www.freepik.com
- Wood vector created by terdpongvector - www.freepik.com
- Business vector created by makyzz - www.freepik.com
- Vintage vector created by freepik - www.freepik.com
- Design vector created by freepik - www.freepik.com
- Party vector created by freepik - www.freepik.com
- Food vector created by freepik - www.freepik.com
- Business vector created by rawpixel.com - www.freepik.com

References

- Binsted, K., & Ritchie, G. (1994). An implemented model of punning riddles. In proceedings of AAAI-94.
- Ritchie, G., Manurung, R., Pain, H., Waller, A., & O'Mara, D. (2006). The standup interactive riddle-builder. IEEE Intelligent Systems, 21(2), 67-69.
- Manurung, R., Ritchie, G., Pain, H., Waller, A., O'Mara, D., & Black, R. (2008). THE CONSTRUCTION OF A PUN GENERATOR FOR LANGUAGE SKILLS DEVELOPMENT. Applied Artificial Intelligence, 22(9), 841-869.
- Román, Iván Guerrero et al. "TheRiddlerBot: A next step on the ladder towards creative Twitter bots." ICCC (2015).