Slides: thomaswinters.be/clin

TorfsBot or Not?

Evaluating User Perception on Imitative Text Generation

by Thomas Winters
@thomas_wint







Rik Torfs

Professor Canon Law & Former Rector KU Leuven

Rik Torfs
@torfsrik

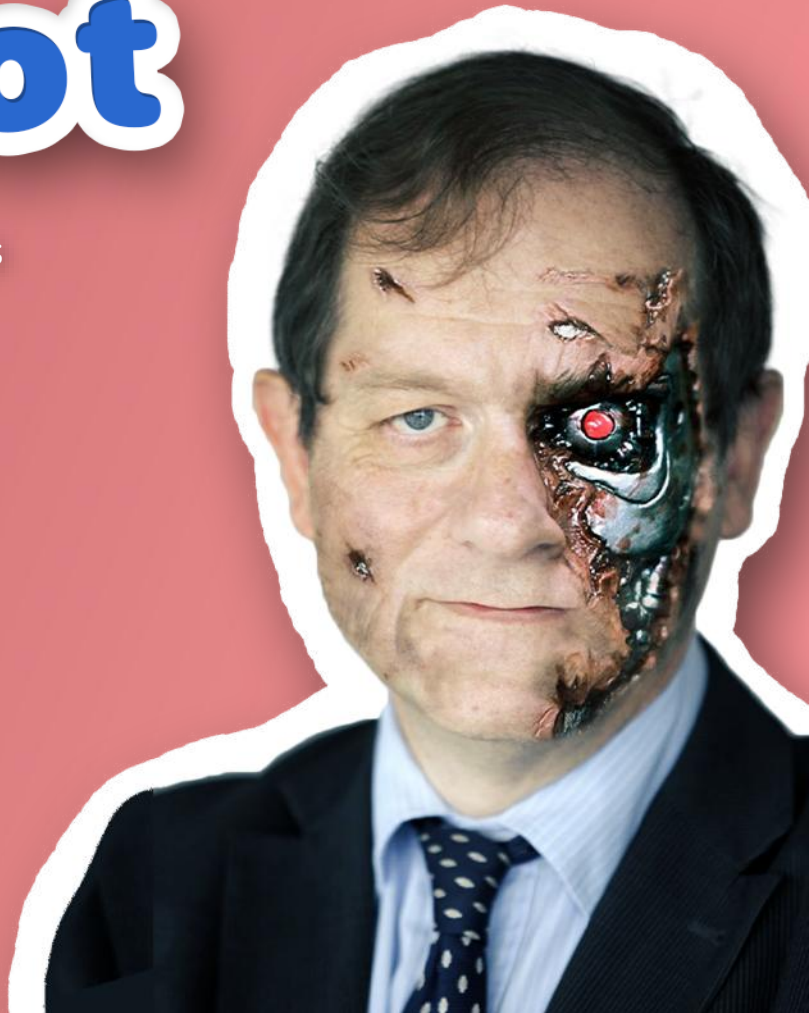
Het mooie van God is dat hij het iedereen van harte gunt niet in hem te geloven.

Rik Torfs
@torfsrik

Redelijkheid wordt stilaan een gevaarlijke vorm van extremisme.

TorfsBot

Fully-automated Twitterbot imitating Rik Torfs
Trained on his tweets & columns
Launched in 2016



Robot Torfs
@TorfsBot

Beste, gevolgd door een komma.
Mails om niet te herinneren.

Robot Torfs
@TorfsBot

Zonder twijfel zien we beter wat er niet is.

TorfsBot

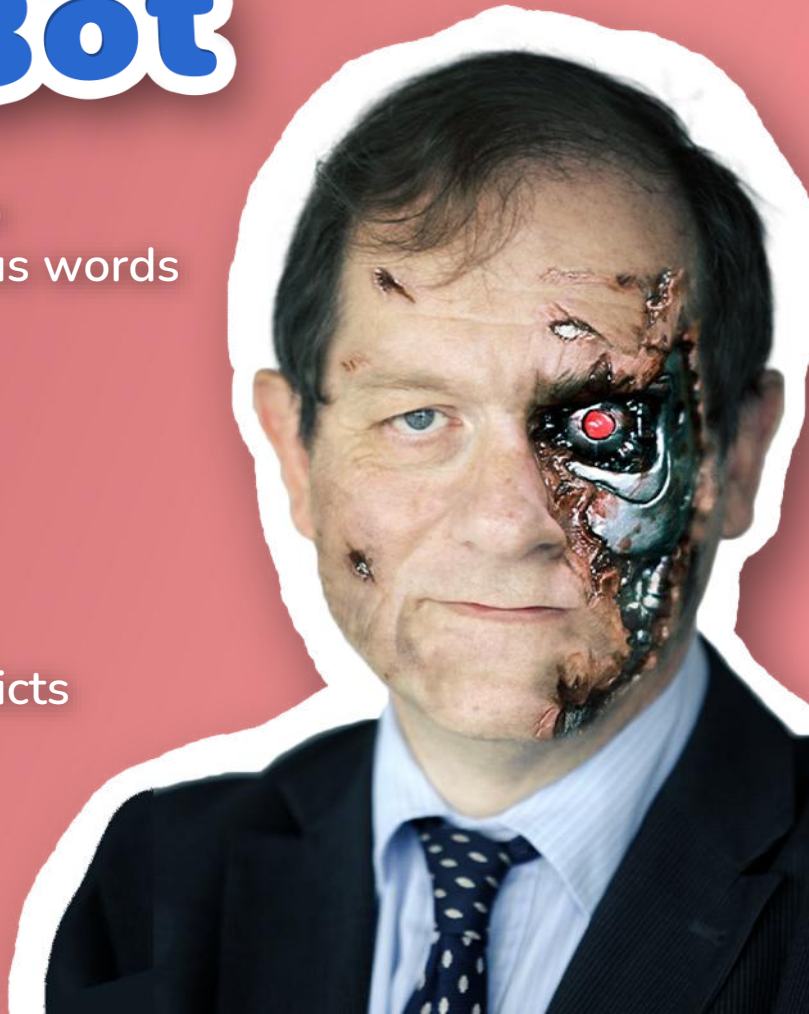
Algorithm 1: Markov Chain (Interpolated 3→5-gram)

1. Counts frequency of word after previous words in tweets & columns van Rik Torfs



2. Takes start words and repeatedly predicts

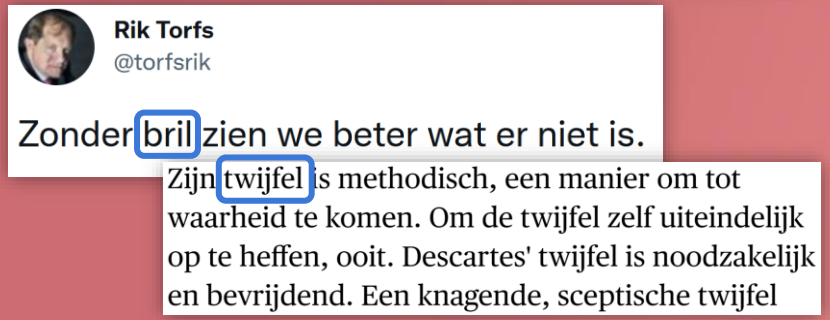




TorfsBot

Algorithm 2: Dynamic template

1. Take random tweet and a random column lines



2. Change tweet key words with words from column with same part-of-speech





Local cohesion > Global form

Markov chain 35% more interactions than dynamic template

(17K interactions on 8K tweets over from June 2017 to June 2023)

Previous work: *Winters, T. (2019).Generating Philosophical Statements using Interpolated Markov Models and Dynamic Templates. 31st European Summer School in Logic, Language and Information Student Session Proceedings, 181–189.*

#Interactions = Quality Parody?

Believable parody?

Funny fail?





Tweets daily poll with random tweet from either Rik Torfs or TorfsBot



Results

602 polls

47K votes total *(avg 79/poll)*

71% correct votes

87% correct majority

68 TorfsBot success

12 Rik Torfs fails

Is de tweet van Rik Torfs of TorfsBot?
@TorfsBotOrNot

Follow

Wat is waarheid? Alles relativeren is al te absoluut.

Rik Torfs	33.7%
TorfsBot	**66.3%**

101 votes · Final results

10:00 AM · May 2, 2022

1 1

Post your reply

Reply

Is de tweet van Rik Torfs of TorfsBot? @TorfsBotOrNot · May 2, 2022

❌ Het is helaas een échte tweet van Rik Torfs...

> **Rik Torfs** @torfsrik · Aug 9, 2014
> Wat is waarheid? Alles relativeren is al te absoluut.

3 4 22

Is de tweet van Rik Torfs of TorfsBot?
@TorfsBotOrNot

Follow

Niemand heeft recht op geluk. Maar zelfoverschatting staat het wel vaker in de weg.

Translate post

Rik Torfs	34.3%
TorfsBot	**65.7%**

67 votes · Final results

10:00 AM · Aug 5, 2022

1 3

Post your reply

Reply

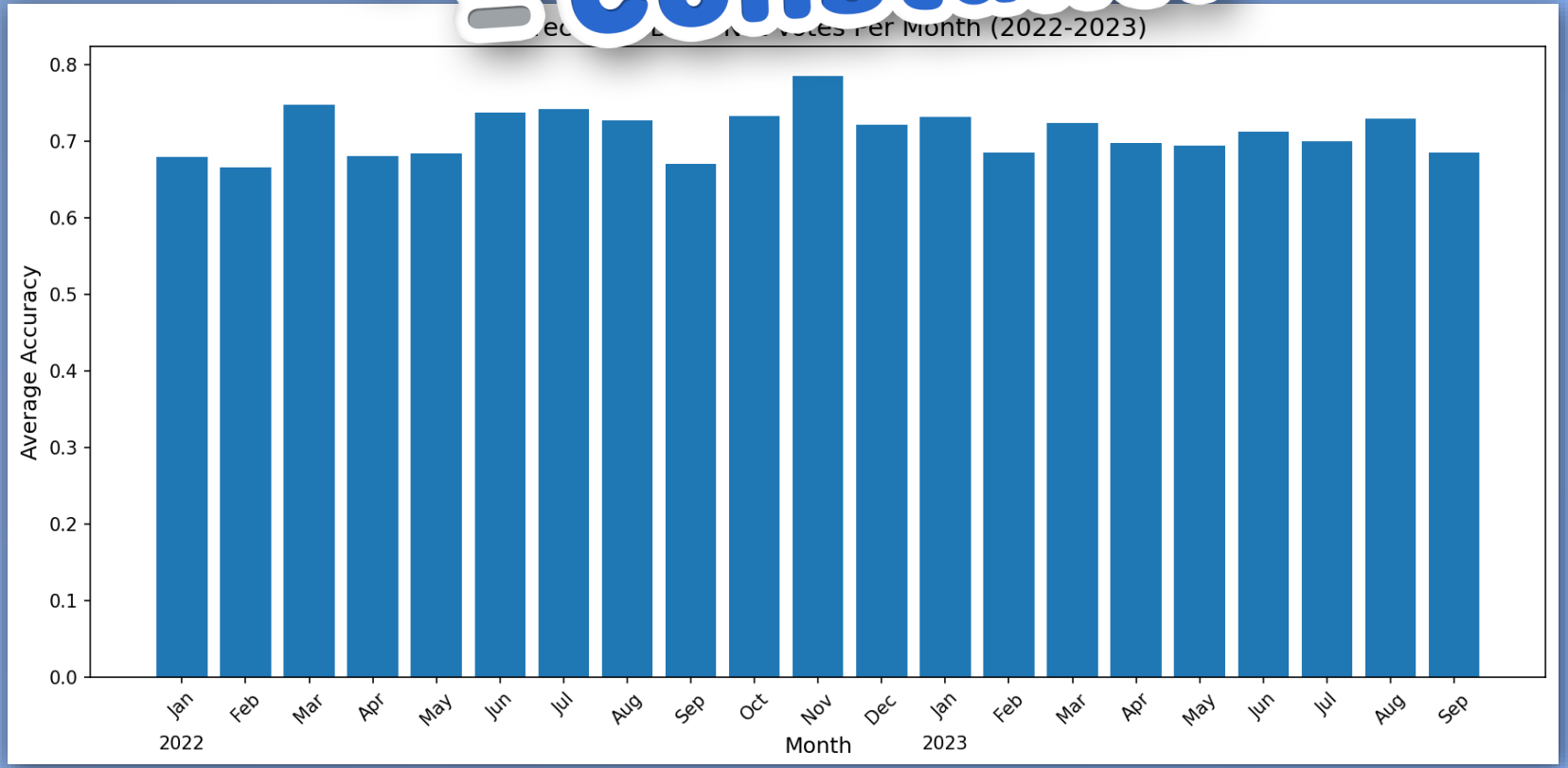
Is de tweet van Rik Torfs of TorfsBot? @TorfsBotOrNot · Aug 5, 2022

❌ Het is helaas een échte tweet van Rik Torfs...

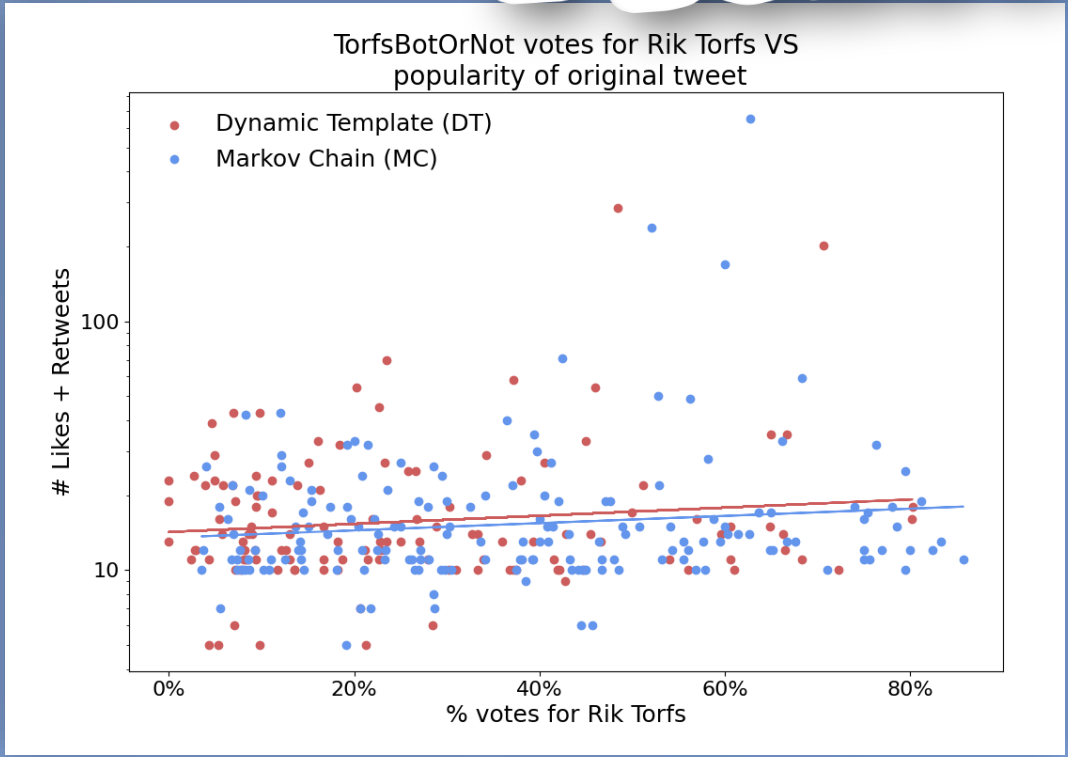
> **Rik Torfs** @torfsrik · Aug 14, 2021
> Niemand heeft recht op geluk. Maar zelfoverschatting staat het wel vaker in de weg.

1 2 13

Correctness = Constant?



Interactions = Believability?

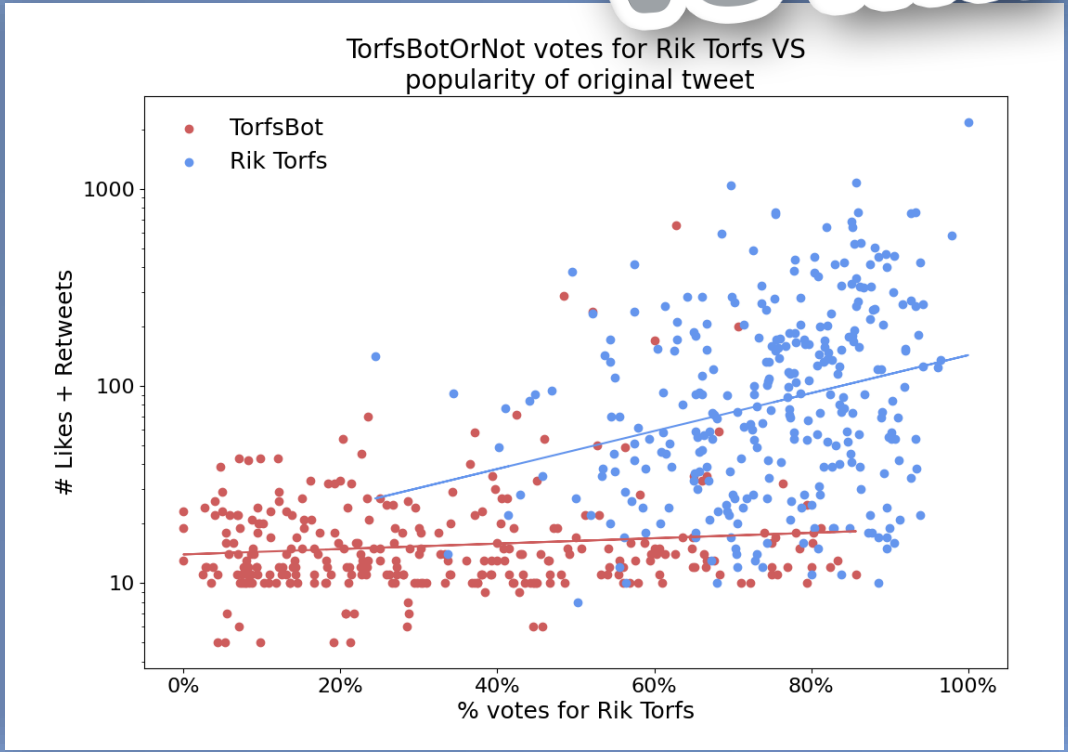


Slightly positively correlated

(0.11 for log(interactions) with %votes)

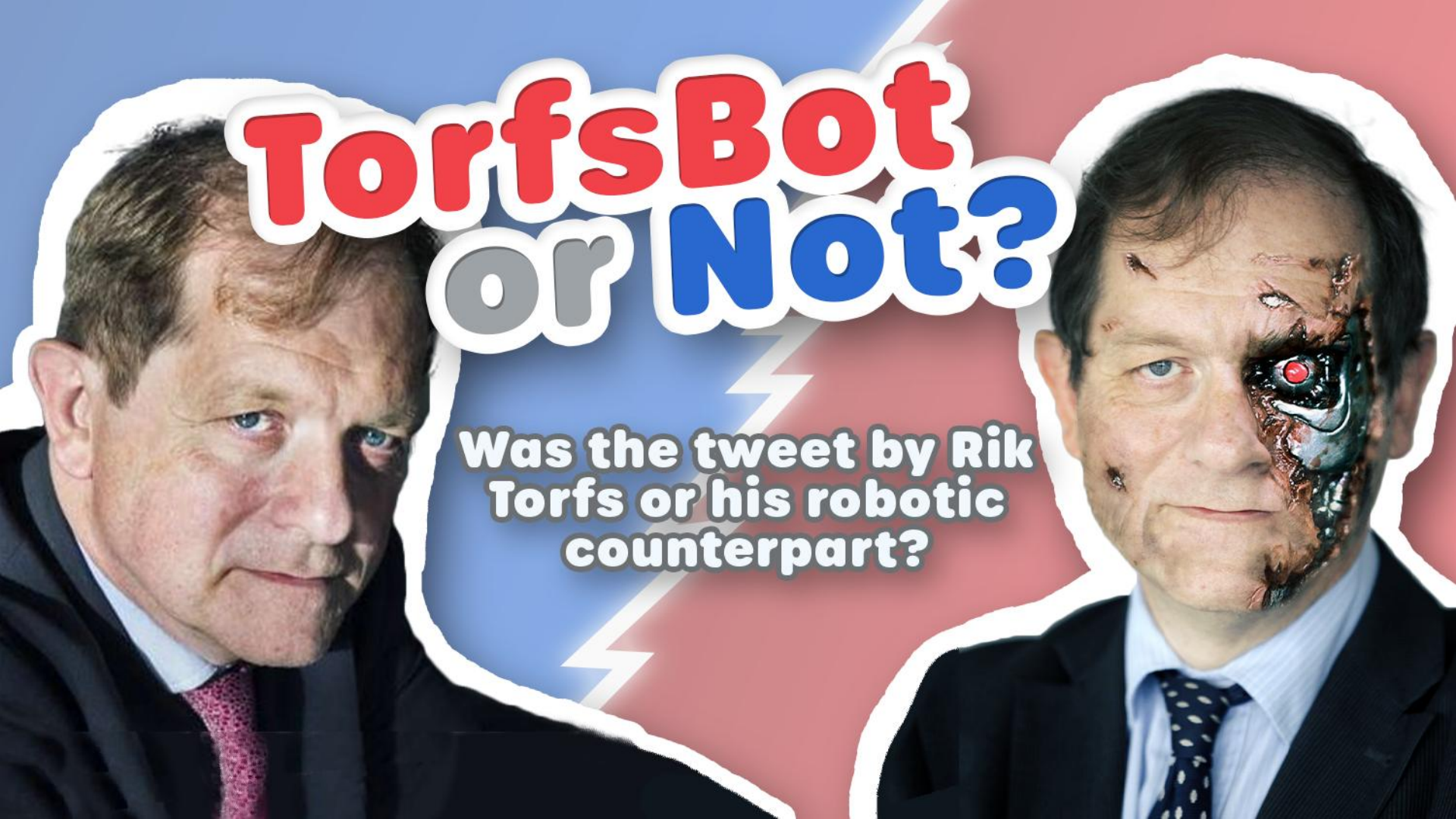
Little difference between algorithms

Interactions VS Rik Votes



More interactions
~ more believable

Weirdly more so
for Rik than bot


TorfsBot
or Not?
Was the tweet by Rik Torfs or his robotic counterpart?

TorfsBot or Not? #1

Als iemand je om je mening bespot, is ze wellicht relevant.

Rik Torfs



Rik Torfs
@torfsrik

Als iemand je om je mening bespot, is ze wellicht relevant.

5:10 p.m. · 28 okt. 2018 · Twitter for iPhone

tweets **4** Geciteerde Tweets **157** Vind-ik-leuks

TorfsBot or Not? #2

Ik wou dat ik twee ideologieën was dan konden we samen terugkeren.

TorfsBot



Robot Torfs
@TorfsBot

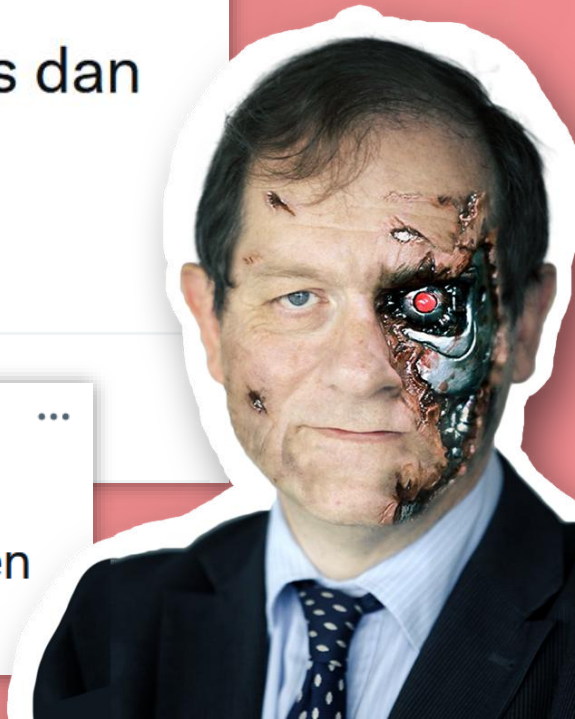
Ik wou dat ik twee ideologieën was dan konden we samen terugkeren.

3:48 p.m. · 27 nov. 2020 · TorfsBot



Rik Torfs
@torfsrik

Ik wou dat ik twee pausen was dan konden we samen spelen.



TorfsBot or Not? #3

Ik geloof in de kracht van radijsjes op recepties.

Rik Torfs





Rik Torfs
@torfsrik

Ik geloof in de kracht van radijsjes op recepties.

7:09 p.m. · 23 mei 2017 · Twitter for iPhone

etweets **4** Geciteerde Tweets **191** Vind-ik-leuks

Rik's Radijsjes

Rik Torfs
@torfsrik

Nu kies ik nog niet voor het grote geld. Maar later ga ik radijsjes kweken.

8:37 p.m. · 12 jan. 2021 · Twitter for iPhone

2 Retweets **1** Citeer Tweet **92** Vind-ik-leuks

Rik Torfs
@torfsrik

"De planeet zit in de puree." Laten we haar redden en enkel nog radijsjes eten.

3:27 p.m. · 12 mei 2018 · Twitter for iPhone

4 Geciteerde Tweets **46** Vind-ik-leuks

Rik Torfs
@torfsrik

Ik geloof in de kracht van radijsjes op recepties.

7:09 p.m. · 23 mei 2017 · Twitter for iPhone

46 Retweets **4** Geciteerde Tweets **191** Vind-ik-leuks

Rik Torfs
@torfsrik

Als we minder nutteloze verplaatsingen maken, hebben we minder auto's nodig, en minder openbaar vervoer. En kunnen we onze radijsjes beter verzorgen.

9:33 p.m. · 4 jun. 2020 · Twitter for iPh

6 Retweets **118** Vind-ik-leuks



TorfsBot or Not? #4

Ideologie troost. Bijvoorbeeld de ideologie dat wij maar erg weinig weten.

TorfsBot



Robot Torfs
@TorfsBot

Ideologie troost. Bijvoorbeeld de ideologie dat wij maar erg weinig weten.

6:39 p.m. · 10 aug. 2021 · TorfsBot

5 Retweets **3** Geciteerde Tweets **13** Vind-ik-leuks



TorfsBot or Not? #5

De laatste TorfsBot Or Not!

Wat je loslaat, lost vaak zichzelf op.

Bond zonder Naam



Bond zonder Naam

Slides: thomaswinters.be/clin

TorfsBot or Not?

Evaluating User Perception on Imitative Text Generation

by Thomas Winters
@thomas_wint





Robot Torfs
@TorfsBot

Verklaar mij gek, maar ik heb wel gesproken.