University of Antwerp
Faculty of Arts

Automated detection of political trends and vaccine skepticism from social media text

VAIA Presents: Pieter Fivez

Antwerp Text Mining Centre (TEXTUA)

31 May 2022

1. About TEXTUA

About TEXTUA: Antwerp Text Mining Centre

- **Core facility** at the University of Antwerp: provides consultancy and services for text mining

- **Team**:

    - Pieter Fivez (coordinator): computational linguistics
    - Walter Daelemans (spokesperson): computational linguistics
    - Mike Kestemont: digital humanities
    - David Martens: applied data mining
    - José Oramas: explainable artificial intelligence
    - Jef Peeraer: software developer

- **Text mining**: automatic processing and analysis of textual data to extract information

- **Information**:

    - factual (concepts, events, relations)
    - subjective (opinions, sentiment)
    - properties of the author (personality, age etc.)

About CLiPS

- **Research center** at the University of Antwerp:
    - **computational linguistics**
    - (developmental) psycholinguistics
    - sociolinguistics
- **Broad range** of research topics in computational linguistics**:**
    - Automated processing of clinical text
    - Development of conversational agents (e.g. chatbots)
    - Authorship verification and attribution
    - Hate speech detection
- **Prominent recent topic**: automated monitoring of public discourse on social media platforms
    - political issues
    - vaccine skepticism and hesitancy

2. Political trends

Tim Kreutz & Walter Daelemans
(based on earlier work by Tom De Smedt)

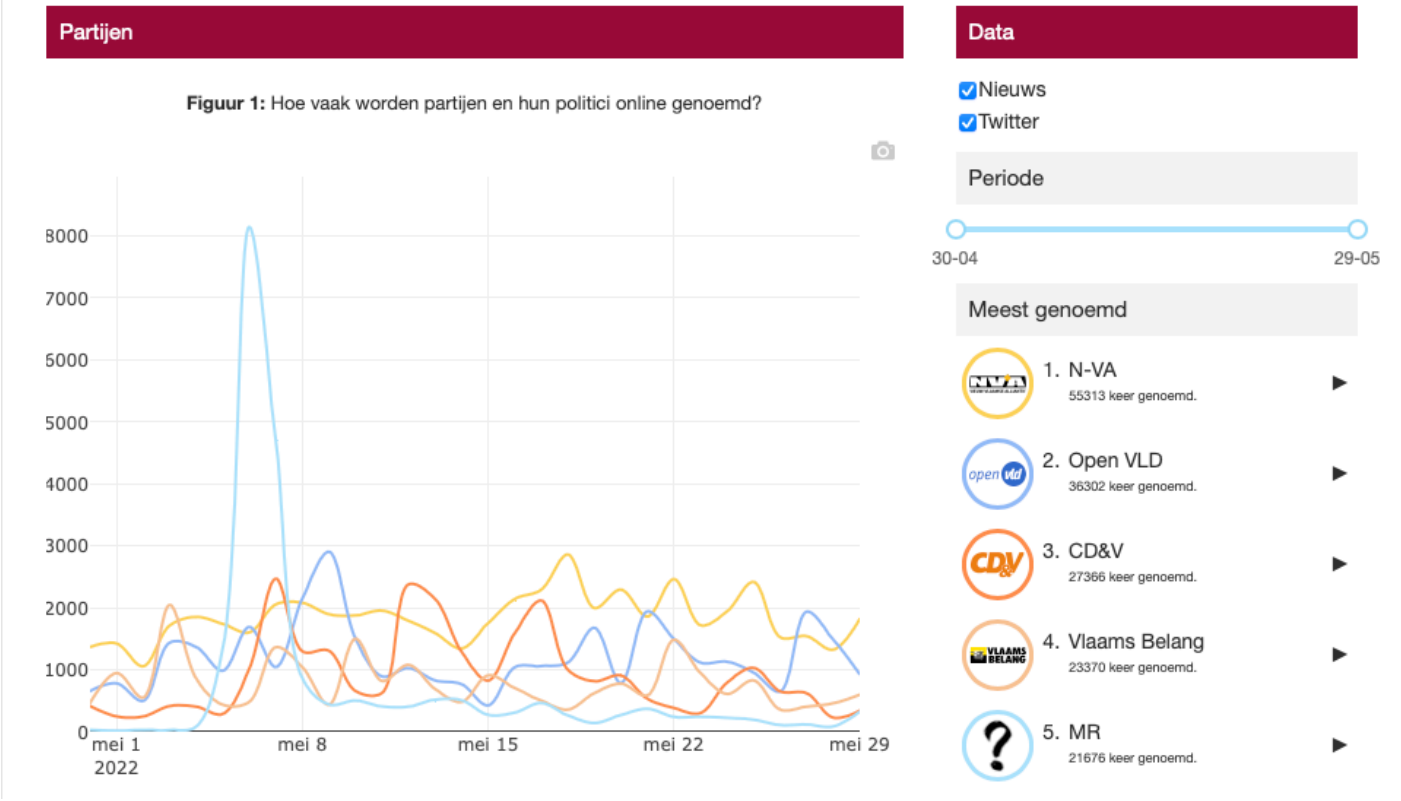
Politieke Barometer

- Part of the **NWS Data** project: interdisciplinary cooperation involving
    - computational linguistics (CLiPS)
    - data mining (Applied Data Mining)
    - political science (M^2P: Media, Movements and Politics)
    - communication science (AMSoC: Antwerp Media in Society Centre)

- **Focus**: what is being written about political parties on social media?
- ± 1000 **news articles** per day (De Standaard, De Morgen, Knack, Humo, …)
- ± 200.000 **tweets** per day

politiekebarometer.uantwerpen.be

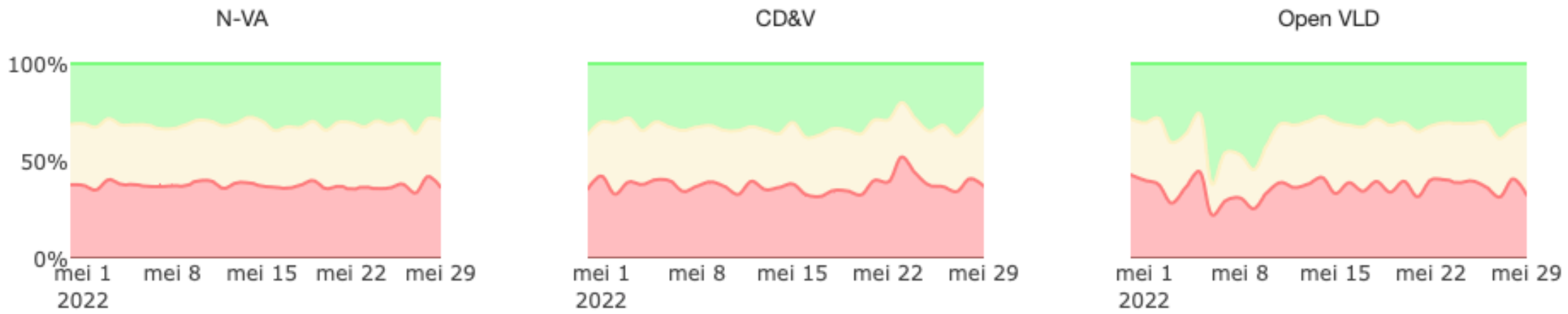
- Mentions of **political parties** and their politicians:



politiekebarometer.uantwerpen.be

- Sentiment towards **political parties:**

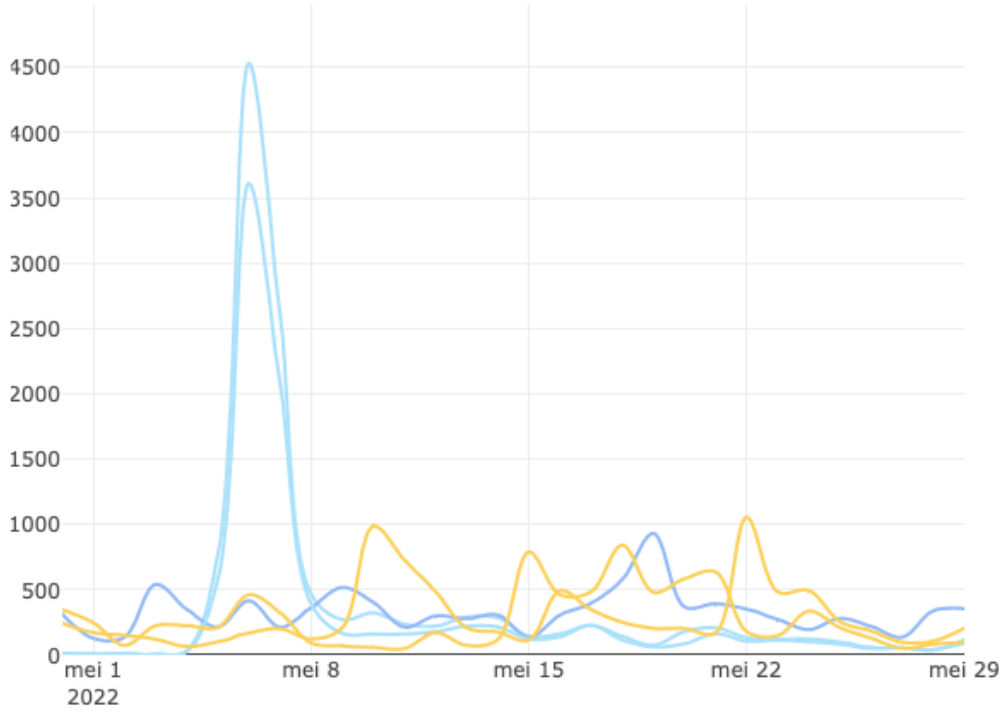
Figuur 2: Hoe evolueert de opinie over partijen op Twitter?

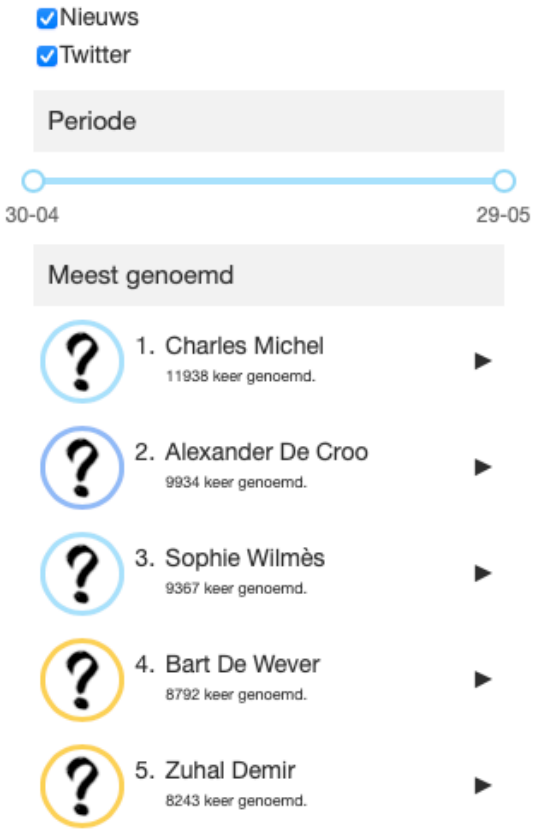


politiekebarometer.uantwerpen.be

- Mentions of **specific politicians:**

Figuur 1: Hoe vaak worden politici online genoemd?

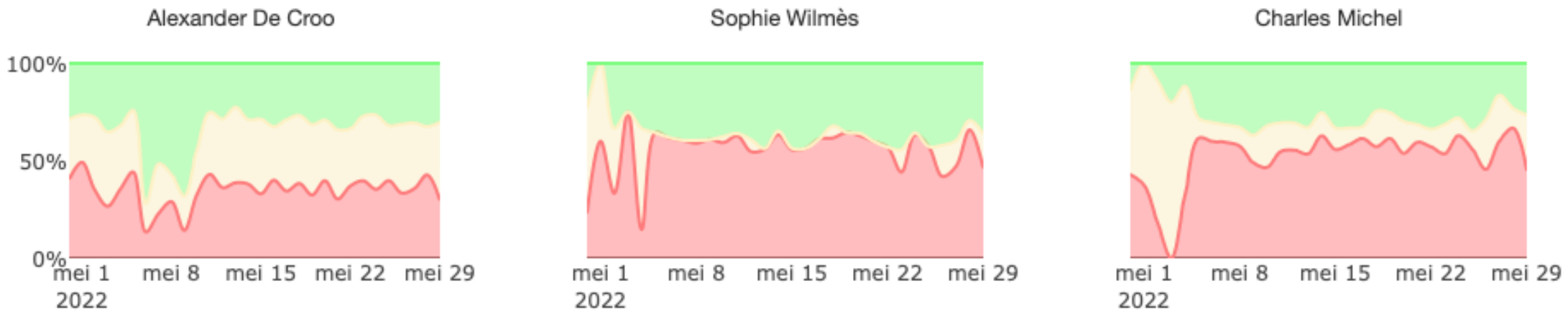




politiekebarometer.uantwerpen.be

- Sentiment toward **specific politicians**:

Figuur 2: Hoe evolueert de opinie over politici op Twitter?



Scientific backbone: relevant ongoing research

- Adapting **sentiment detection** techniques for **specific domains** (e.g. political text)
- Detecting and **classifying political issues** in different text genres:
    - social media texts
    - news articles
    - political party manifestos
    - parliamentary documents

Related scientific publications from CLiPS

- Sentiment detection: https://aclanthology.org/C18-1090/
- Detecting political issues: https://www.clips.uantwerpen.be/~walter/papers/2021/kreutz+2021.pdf
- Twitter search: https://www.clips.uantwerpen.be/clinjournal/clinj/article/view/92
- More related publications: https://scholar.google.be/citations?user=LVpMl34AAAAJ&hl=nl

3. Vaccine skepticism

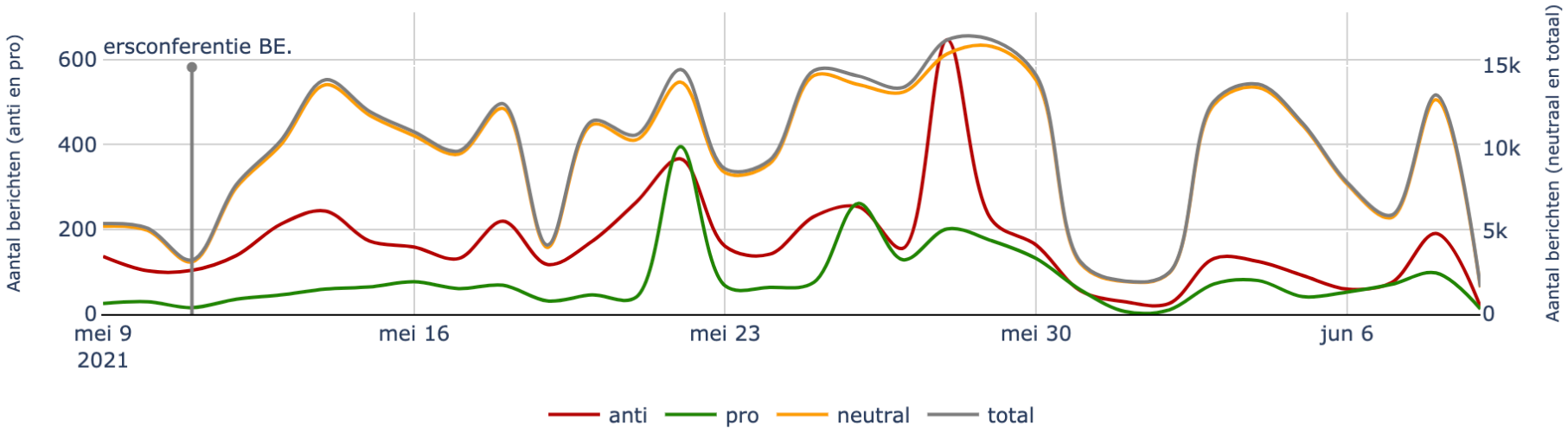
3.1. Vaccinpraat

Jens Lemmens, Jens Van Nooten, Tim Kreutz & Walter Daelemans
(demographic predictions provided by Textgain)

vaccinpraat.be

- **Sentiment** about COVID-vaccination on Twitter / Facebook

Figuur 1: Hoeveel vaccinatieberichten verschijnen er dagelijks op Twitter?



vaccinpraat.be

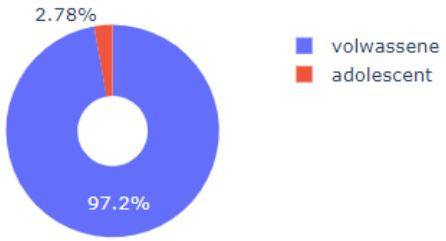
- Where are the **hotspots** for vaccine hesitancy?

Figuur 1: Waar komen de antivaccinatieberichten vandaan?

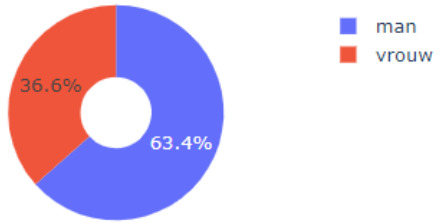
vaccinpraat.be

- What is the **profile of the authors?**

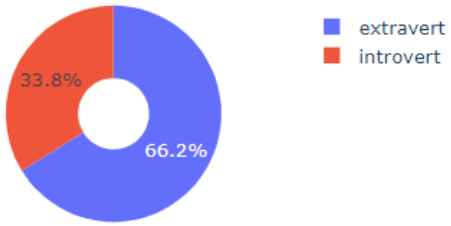
Figuur 2: Welke leeftijd hebben auteurs van antivaccinatieberichten?



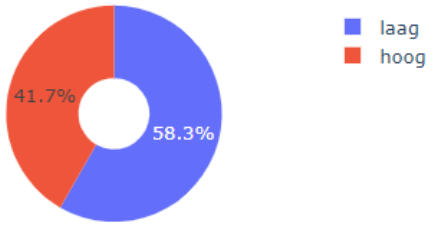
Figuur 3: Welk geslacht hebben auteurs van antivaccinatieberichten?



Figuur 4: Welke persoonlijkheid hebben auteurs van antivaccinatieberichten?



Figuur 5: Welk opleidingsniveau hebben auteurs van antivaccinatieberichten?



vaccinpraat.be

- What do they **talk about?**

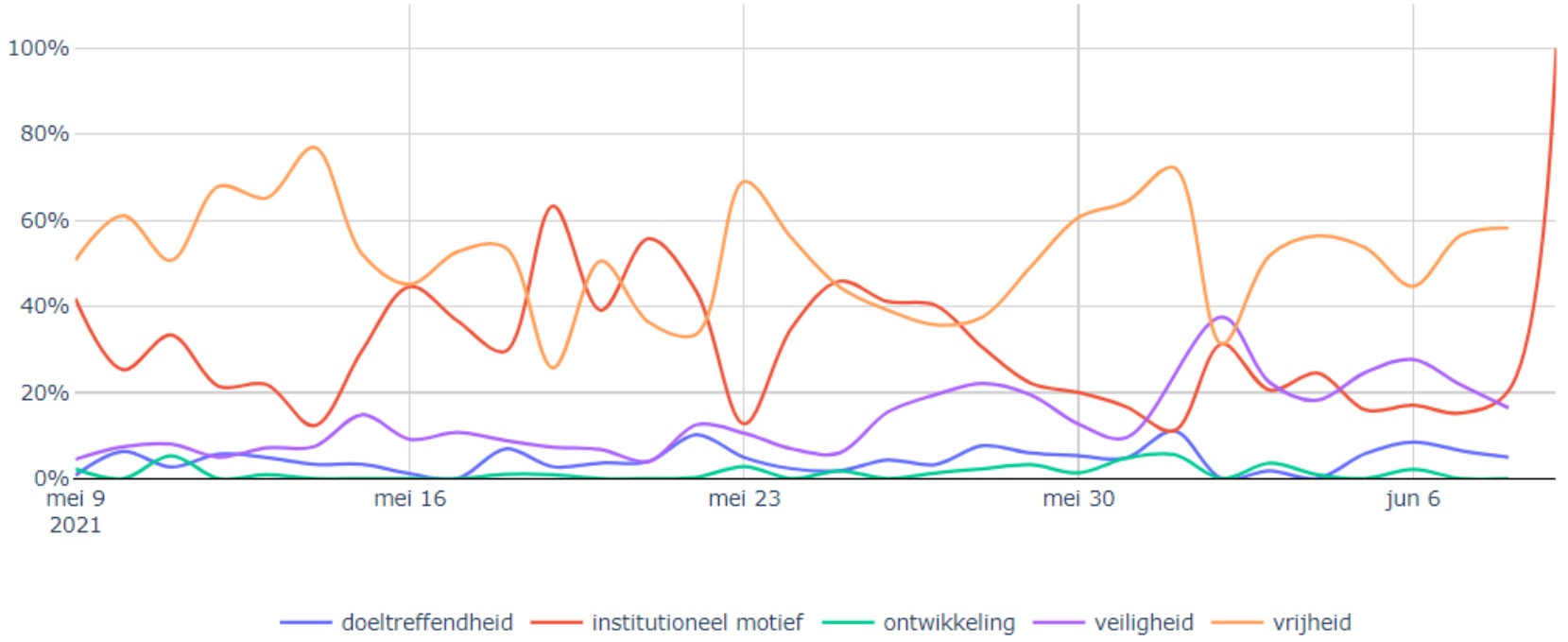
Figuur 1: Welke personen (grijs), locaties (blauw) en organisaties (rood) worden het vaakst genoemd in antivaccinatieberichten?

Figuur 2: Welke hashtags worden het vaakst gebruikt in antivaccinatieberichten?

vaccinpraat.be

- What **narratives & motivations** used by anti-vax authors?

Figuur 3: Welke motivaties komen het vaakst voor in antivaccinatieberichten?



Scientific backbone: relevant ongoing research

- **Domain adaptation**: adapting pre-trained language models for vaccination discussions on social media
    - better abstract representations of the nuanced “meaning” of words in this use context
- Identifying **vaccine skepticism** from **heterogeneous information networks** using:
    - tweets
    - user biographies
    - user follows
    - retweets
- Detecting vaccine skepticism at **different levels of categorization**
    - **binary**: skeptic or not
    - **argument detection**: which specific arguments?

Related scientific publications from CLiPS

- Detecting vaccine skepticism: https://clinjournal.org/clinj/article/view/134
- Domain adaptation for vaccination discussions: https://arxiv.org/abs/2203.07362

3.2. Vaccinchat

Jeska Buhmann, Maxime De Bruyn, Ehsan Lotfi & Walter Daelemans

University of Antwerp
Faculty of Arts

vaccinchat.be

- Automated **question-answering** for COVID vaccination
  - uses **official information**
  - hosted in **open source** software environment (RASA)
- Covers **± 150 different answer types**
  - Trained on 50-1000 question-answer pairs per answer type
  - Evaluation: **± 90% accurate**
- Technology: **Natural Language Understanding (NLU)** pipeline
  - Deep Learning with neural networks
  - Natural Language Processing (NLP) with classical machine learning

vaccinchat.be

- Example: **vaccination hesitancy**

Scientific backbone: relevant ongoing research

- Better abstract **representations of the "meaning"** of questions

- Train chatbots indirectly on **FAQ-data and/or collections of relevant documents**

    - benefit: **quick development** of new chatbots

- **Personalisation** of chatbots:

    - **emotion** detection
    - **empathic discourse** for detected emotions

- Adapt conversational agents for **specific tasks**

Related scientific publications from CLiPS

- Response generation: https://arxiv.org/abs/2110.02067
- FAQ answering for vaccination: https://arxiv.org/abs/2108.00719
- Multilingual FAQ answering: https://arxiv.org/abs/2109.12870
- Domain adaptation for COVID: http://ceur-ws.org/Vol-2666/KDD_Converse20_paper_7.pdf

4. Practical information

Practical information

- www.uantwerpen.be/textua
- pieter.fivez@uantwerpen.be

- **Book a meeting**

    - **free** 1-hour in-depth meeting to analyse case
    - we devise a plan of action

- **Pricing**

    - **accessible pricing**: we want to create a diverse portfolio
    - **routine tasks**: simple transfer of money with an invoice
    - **larger research tasks**: draw up a specific contract

University of Antwerp
Faculty of Arts

Automated detection of political trends and vaccine skepticism from social media text

VAIA Presents: Pieter Fivez

Antwerp Text Mining Centre (TEXTUA)

31 May 2022