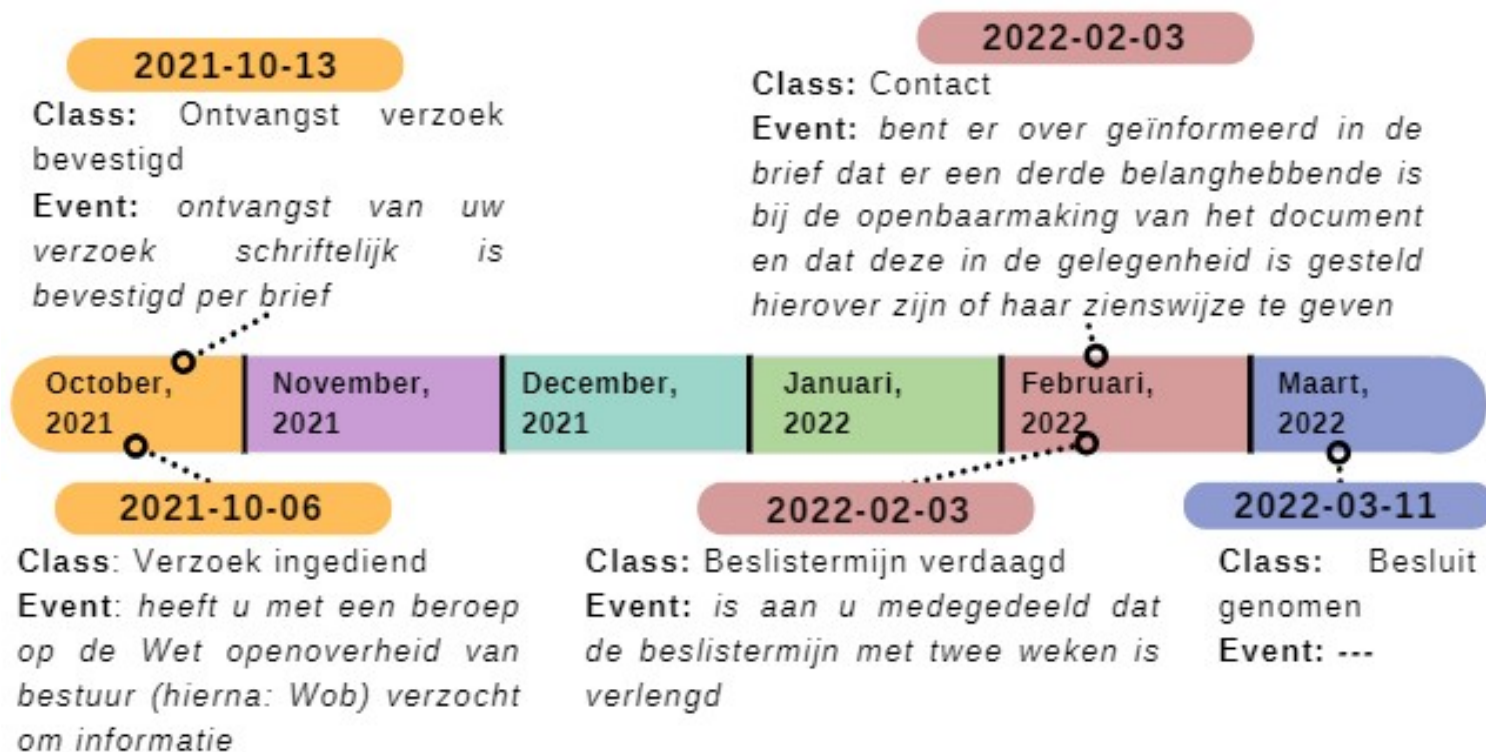


Tijdslijn extractie uit Woo-besluitbrieven met ChatGPT



Femke Bakker

Hoofdvraag

Hoe goed kan ChatGPT automatisch alle datums met hun gebeurtenissen uit een Woo-besluitbrief extraheren en op een tijdlijn plaatsen?

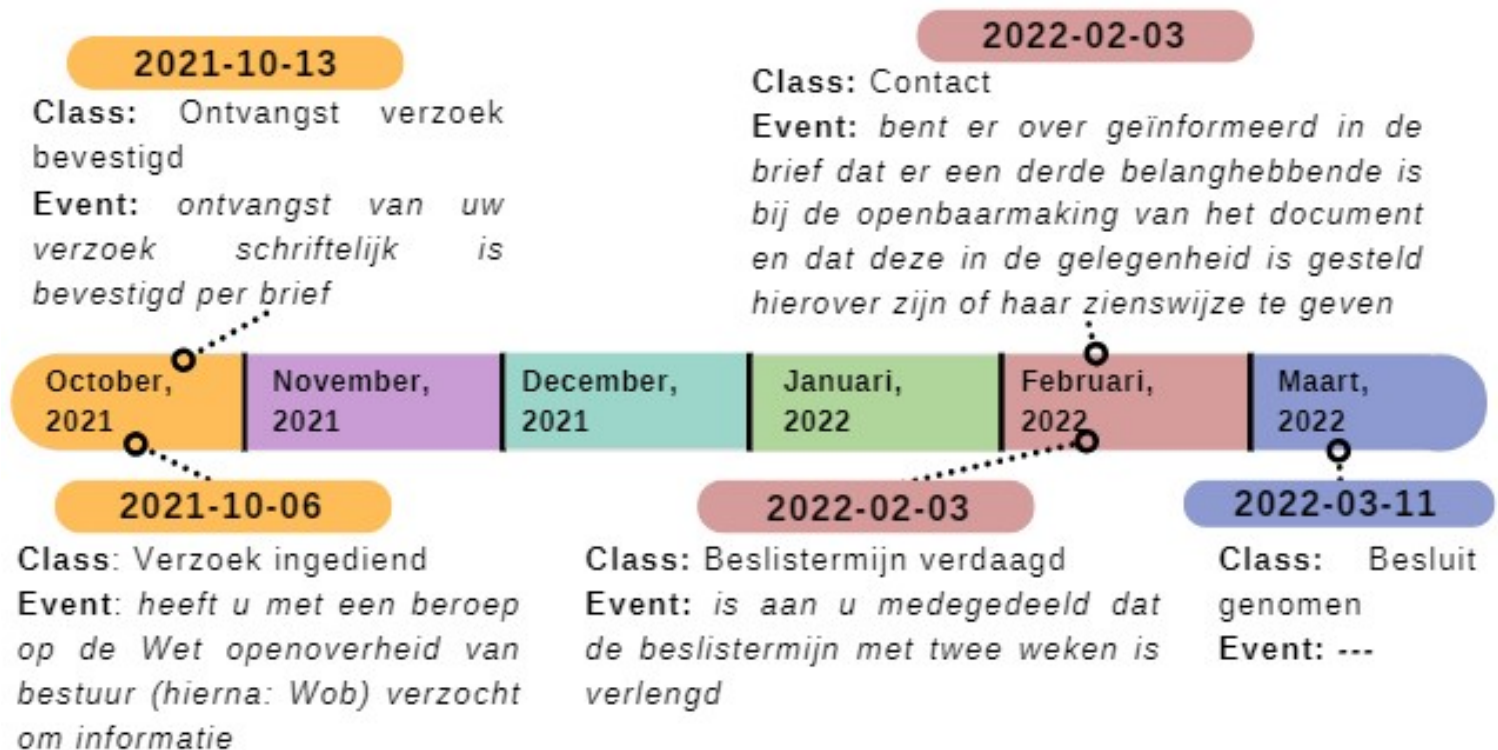
Relevantie

Toepassingen tijdlijnen:

- grafische samenvattingen
- process mining
- process monitoring
- meta data

Tijddlijnen

Event classes: *besluit genomen, verzoek ingediend, verzoek ontvangen, ontvangst verzoek bevestigd, beslistermijn verdaagd, inwerking treden van Woo, contact, overig*



Deeltaken tijdslijn extractie

1. Onderscheid maken tussen datums met gebeurtenis en datums zonder.
2. Gebeurtenis omschrijvingen extraheren.
3. Datum en gebeurtenis matchen.
4. Gebeurtenis classificeren.

Onderdelen van onderzoek

1. Dataset verzameling
2. Ontwikkelen van tijdlijn extractie pipeline
3. Evalueren

Dataset verzameling

- 100 PDF's met 525 datums met een gebeurtenis
- NLP software pakt zinnen met een datum erin
- 4 deeltaken tijdens annoteren:
 - datums corrigeren
 - gebeurtenis omschrijving labelen (als het bestaat)
 - gebeurtenis classificeren
 - relatie tussen datum en gebeurtenis labelen
- 50/50 development/test set

Tijddlijn extractie pipeline

1. Split tekst in zinnen en selecteer datum
 - 1.1. Extraheer besluit datums met Regex.
2. Prompt ChatGPT om het volgende te doen voor elke datum:
 - 2.1. selecteer gebeurtenis omschrijving
 - 2.2. classificeer gebeurtenis (als gebeurtenis gevonden is)
3. Sla datum, gebeurtenis omschrijving en gebeurtenis classe combinaties op. Sorteer ze chronologisch en maak een tijddlijn.

Resultaten

- Step 1: Accuracy datum extractie en correctie = 98% (N=259)
- Step 1.1: Accuracy besluit datum = 99% (N=541)

4 taken ChatGPT

- Step 2.1: Accuracy classificeren datum met/zonder gebeurtenis = 54% (N=493)
- Step 2.1: Jaccard Similarity gebeurtenis omschrijving = 50.2% (N=144)
- Step 2.1: Accuracy datum-gebeurtenis matching = 79% (N=67)
- Step 2.2: Accuracy scores gebeurtenis classificatie = 80% (N=205)

Complete tijdlijnen

- Fouten: gemiddelde=5.6 & modus=4
- 8% van tijdlijnen compleet foutloos
- Kosten: 0.02 USA dollar
- Tijd: 1 minuut & 2 sec

Conclusie

Deeltaken van ChatGPT

- verbetering aangeraden:
 - classificeren datum met/zonder gebeurtenis (accuracy=54%)
 - extraheren gebeurtenis omschrijving (Jaccard Similarity=50.2%)
- werken goed:
 - datum-gebeurtenis matching (accuracy=79%)
 - classificeren gebeurtenis (accuracy=80%)

Hoofdvraag

- ChatGPT kan redelijk accuraat de tijdlijnen extraheren
- Verder onderzoek: Vergelijken met QA model [Du & Cardie, 2020]

Bedankt!

Zijn er nog vragen?