

# Generative AI Tools for Education

Walter Daelemans

[www.uantwerpen.be/textua](http://www.uantwerpen.be/textua)

## TEXTUA

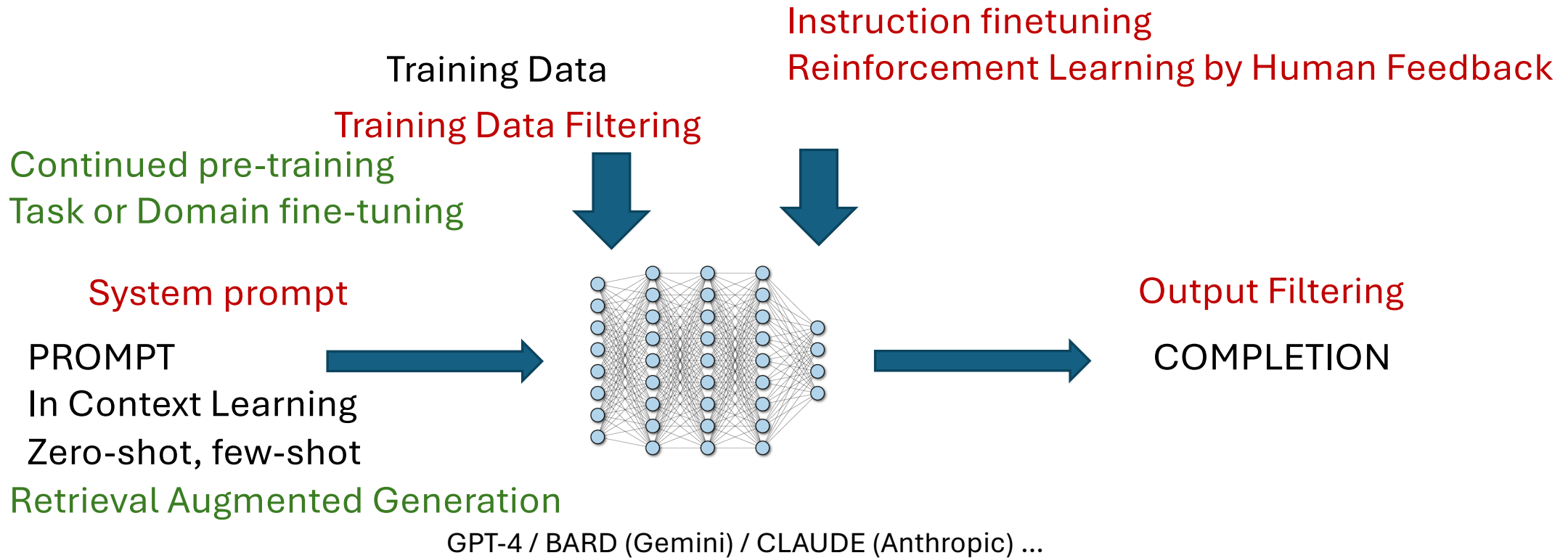
Antwerp Text Mining Centre



# Overview

- Recent developments in Large Language Models (LLMs):
  - GPT Assistants improve usability
  - Language-specific LLMs (GPT-NL)
  - <https://www.tno.nl/nl/digitaal/digitale-innovatie/data-sharing/generative-ai/gpt-nl-versterkt-nederlandse-autonomie/>
  - Retrieval Augmented Generation (RAG) avoids hallucinations
  - Open Source Models for flexibility, privacy, security and low cost
- Three projects I'm working on in TEXTUA context:
  - An assistant for ESL essay feedback / grading
  - Large Language Model output detector for Dutch academic text
  - (UAntwerp house style expert)

# Alignment Adaptation



<https://chat.lmsys.org/?leaderboard>

Rank ▲	🤖 Model ▲	★ Arena Elo ▲	📊 95% CI ▲	🗳️ Votes ▲	Organization ▲	License ▲	Knowledge Cutoff ▲
1	<a href="#">GPT-4-Turbo-2024-04-09</a>	1259	+5/-5	21448	OpenAI	Proprietary	2023/12
1	<a href="#">GPT-4-1106-preview</a>	1254	+4/-4	66858	OpenAI	Proprietary	2023/4
1	<a href="#">Claude 3 Opus</a>	1253	+3/-3	66357	Anthropic	Proprietary	2023/8
2	<a href="#">GPT-4-0125-preview</a>	1249	+3/-4	54002	OpenAI	Proprietary	2023/12
5	<a href="#">Bard (Gemini Pro)</a>	1209	+6/-6	12435	Google	Proprietary	Online
5	<a href="#">Claude 3 Sonnet</a>	1202	+3/-3	69817	Anthropic	Proprietary	2023/8
5	<a href="#">Llama-3-70b-Instruct</a>	1198	+9/-11	2754	Meta	Llama 3 Community	2023/12
7	<a href="#">Command R+</a>	1193	+3/-3	37752	Cohere	CC-BY-NC-4.0	2024/3
7	<a href="#">GPT-4-0314</a>	1189	+4/-4	45288	OpenAI	Proprietary	2021/9
10	<a href="#">Claude 3 Haiku</a>	1181	+3/-3	61623	Anthropic	Proprietary	2023/8
11	<a href="#">GPT-4-0613</a>	1165	+3/-3	63963	OpenAI	Proprietary	2021/9
12	<a href="#">Mistral-Large-2402</a>	1158	+3/-4	40831	Mistral	Proprietary	Unknown
12	<a href="#">Qwen1.5-72B-Chat</a>	1153	+3/-4	30800	Alibaba	Qianwen LICENSE	2024/2
12	<a href="#">Claude-1</a>	1150	+5/-4	21824	Anthropic	Proprietary	Unknown
12	<a href="#">Mixtral-8x22b-Instruct-v0.1</a>	1147	+9/-11	2945	Mistral	Apache 2.0	2024/4

# TEXTUA GPT Forge

- Embracing the Open Source models
- Dedicated H100 GPU 80 Gb (+ CLiPS server + CALCUA + cloud)
- OS model (Mistral or Meta or Cohere)
- Software for *our own* instruction fine-tuning, continued pre-training, task and domain fine-tuning, RAG, ...
- GPT Assistants available inside secure UAntwerp network

# Essay feedback and grading

<https://chat.openai.com/g/g-RCnBHK6Ve-essay-feedback>

- GPT4! (Paying subscription)
- Can be made with Open AI API
  - More flexible
- Can be made through Open AI editor
  - Less flexible, easier
- Good prompt engineering still important

Create

Configure

### Name

essay feedback

### Description

provide feedback on an essay

### Instructions

You are an expert teacher of English. You will give structured feedback on a draft student essay that will be input taking into account a number of criteria and giving concrete examples of what can be improved in the essay for each of the criteria.

Do this step by step:

1. Read the criteria which are listed between <criteria> markers below.
2. Read the draft student essay which will be input.
3. Provide for each criterion

### Conversation starters

Please enter an essay.



### Knowledge

If you upload files under Knowledge, conversations with your GPT may include file contents. Files can be downloaded when Code Interpreter is enabled

Upload files

### Capabilities

- Web Browsing
- DALL-E Image Generation
- Code Interpreter [?](#)



# Prompt for feedback on ESL essays

You are an expert teacher of English. You will give structured feedback on a draft student essay, taking into account a number of criteria and giving concrete examples of what can be improved in the essay for each of the criteria.

Do this step by step:

1. Read the criteria which are listed between <criteria> markers below.
2. Read the draft student essay which is between the <essay> markers below.
3. Provide for each criterion
  - a) a SHORT general appreciation for that criterion
  - b) Suggestions for change: QUOTE parts TAKEN FROM THE ESSAY that should be improved for that criterion and provide a suggested correction, or if there aren't any problems for that criterion: "none"

<criteria> {criteria} <criteria>

<essay> {essay} <essay>

# Criteria

(Based on Marilize Pretorius' course)

- Formal and objective register
- Structure
- Argumentation
- Referencing
- Paraphrasing
- Quoting
- Grammar and Lexicon
- Clarity
- Spelling and Punctuation

WA

**You**

What if we could tell how all of the people we were without even asking. Most people can tell when someone is sad, happy, or mad just by looking at them, but what happens when that person is hiding how they are truly feeling. What then? Well new technology can allow us to see exactly how someone, who is trying to hide their emotions, is feeling. This is unfair for many reasons. The technology may be cool but people have a right to keep how they are feeling a secret.

The computer software works by making a 3-D computer model of the face; the computer watches the different muscles in the face move to identify how the person is feeling. The technology developed by Dr. Huang can even calculate mixed emotions with the exact percent of how strong the emotions are. Dr. Huang observes that "facial expressions for each emotion are universal" I thought people expressed emotions in different ways though? Dr. Huang also states that "even though individuals often show varying degrees of expression."

People have a right to not tell people how they are feeling. Did we really need to know the Mona Lisa was 9 percent disgusted? Now that raises the question of why? Was Leonardo de Vinci's hair messed up? Or did she think he needed to take a shower? What is the point in knowing someone is feeling without knowing why? We do not need technology to ask someone how they are feeling. or even why they are feeling the way they are feeling that way.

There are really no benefits to knowing how someone is feeling without knowing why. The technology is very cool. It is crazy how Dr. Huang was able to develop a computer that uses muscle movement to identify emotions. The technology is useless though unless you are prepared and willing to ask why?



**essay feedback**

Thank you for sharing your essay draft. I will provide feedback based on the listed criteria:



## 1. Language Used

- **Appreciation:** The language is somewhat informal for an academic context.
- **Suggestions:** Consider revising colloquial expressions and informal phrasing to match an academic tone. For example:
  - "What if we could tell how all of the people us were without even asking" (line 1) could be revised to "What if we could understand everyone's emotional state without inquiry?"
  - "The technology may be cool but people have a right to keep how they are feeling a secret" (line 4) could be revised to "While the technology is innovative, individuals maintain the right to conceal their emotions."

## 2. Introduction

- **Appreciation:** You introduce the topic, but it's not clearly outlined for the reader.
- **Suggestions:** Provide clearer background information and state the essay's aim. Modify line 1 to something like: "This essay explores emerging technology that interprets hidden emotions and discusses the ethical implications of such capabilities."

## 3. Paragraph Structure

- **Appreciation:** Some paragraphs lack a clear central idea.
- **Suggestions:** Ensure each paragraph centers around a single, coherent idea that supports your thesis. For instance, the first paragraph mixes the introduction of technology with ethical concerns; separate these ideas into distinct paragraphs for clarity.

## 11. Grammar

- **Appreciation:** Some grammatical issues distract from readability.
- **Suggestions:** Correct errors such as "there emotions" to "their emotions" (line 3) and "secert" to "secret" (line 4).

## 12. Lexicon

- **Appreciation:** The word choice is occasionally informal or imprecise.
- **Suggestions:** Replace casual or vague terms with more precise academic language, e.g., "cool" could be "innovative" or "advanced."

## 13. Clarity

- **Appreciation:** Some sentences are ambiguous or confusing.
- **Suggestions:** Clarify confusing sentences like "What if we could tell how all of the people us were without even asking" by rephrasing for clarity.

## 14. Accuracy

- **Appreciation:** Some statements are overly general or unsupported.
- **Suggestions:** Provide specific examples or evidence to support claims, such as the practical implications of emotion-detecting technology.

## 15. Spelling and Punctuation

- **Appreciation:** There are several spelling and punctuation errors.
- **Suggestions:** Correct spelling errors like "techonlogy" to "technology" and "delevopted" to "developed." Ensure consistent punctuation throughout.

# Detecting LLM-generated Text

# Real or Fake?

- De oplader werkte eerst prima maar na nog geen jaar de oplader in gebruik te hebben gehad zijn de oplaadconnectoren van het blokje plots afgebroken. Ik heb er hierdoor sterk mijn twijfels bij of dit wel een officieel Apple product is.
- De screenprotector voor de iPhone 11 is een goede investering. De kwaliteit is uitstekend, het beschermt het scherm optimaal tegen krassen en de installatie is kinderspel. Absoluut een aanrader!

# Real or Fake?

- De oplader werkte eerst prima maar na nog geen jaar de oplader in gebruik te hebben gehad zijn de oplaadconnectoren van het blokje plots afgebroken. Ik heb er hierdoor sterk mijn twijfels bij of dit wel een officieel Apple product is. **[HUMAN]**
- De screenprotector voor de iPhone 11 is een goede investering. De kwaliteit is uitstekend, het beschermt het scherm optimaal tegen krassen en de installatie is kinderspel. Absoluut een aanrader! **[LLM]**



# Techniques

- Watermarking
  - Needs the cooperation of the makers of the LLMs
- Detection
  - Ideas from detecting machine translation
  - Ideas from authorship attribution and profiling (stylometry)
    - According too our software, ChatGPT is a 25+, highly educated, extraverted, man 😊
  - Perplexity & burstiness
    - LLM Language is more consistent and predictable
  - Example software:
    - <https://contentatscale.ai/ai-content-detector/> (98% accuracy for English claimed)

# CLIN33 Shared Task (organized by CLiPS)

## English

	Submission_ID	Macro-accuracy	News	Twitter	Reviews	Poetry	Mystery genre	Open-source
0	Hans_van_Halteren	0.85	0.99	0.69	0.82	0.63	0.99	0.96
1	DetecTUM	0.82	0.93	0.69	0.78	0.80	0.78	0.92
2	Elsevier	0.81	0.98	0.65	0.75	0.50	0.99	0.98
3	NLP_MS	0.74	0.87	0.63	0.63	0.65	0.85	0.82

## Dutch

	Submission_ID	Macro-accuracy	News	Twitter	Reviews	Poetry	Mystery genre
0	Elsevier	0.75	0.95	0.70	0.77	0.50	0.84
1	DetecTUM	0.74	0.96	0.74	0.80	0.53	0.67
2	Hans_van_Halteren	0.72	0.97	0.58	0.78	0.53	0.74
3	NLP_MS	0.71	0.90	0.78	0.70	0.56	0.60

# TEXTUA system (Dutch academic register) coming up

- In development (help welcome)
  - Donating student-generated data
- Currently 95-96% for factual academic text
- What if we reach 98-99%?
  - Is that enough to confront students?
  - Do we want to minimize false positives or false negatives?

UAntwerp house style

# Terminology

aanbesteding	tender
aanbevelingsbrief	letter of recommendation
aankoopdienst	Purchasing Office
aanstaande student	prospective student
aanstelling (op een project)	appointment (to a project)
aanstellingsbrief	letter of appointment
aanstellingspercentage	appointment percentage
aanvullend/extralegaal pensioen	superannuation
AAP (Assisterend Academisch Personeel)	Academic Assistant Staff (AAP)
abituriënt (student in het laatste jaar van het secundair onderwijs)	secondary school-leaver
academicus	academic
academiejaar	academic year
academisch jaar	academic year
academisch medewerker (Bijzonder Academisch Personeel, BAP)	research associate (Contract Research Staff, BAP)
...	

# Schrijfwijzer

## Afspraken

Volgende afspraken zijn gebaseerd op de spellingregels van het *Groene Boekje* en de BIN-normen (Belgisch Instituut voor Normalisatie). Het *Groene Boekje* kan je op internet raadplegen: [www.woordenlijst.org](http://www.woordenlijst.org).

## Adressen

Voor de plaatsnaam in adressen gebruik je hoofdletters. Als je een Antwerps district vermeldt, doe je dat in hoofdletters en schrijf je Antwerpen met kleine letters tussen haakjes. Je schrijft geen landcode voor de postcode, volgens de Europese richtlijnen. Het adres schrijf je zo veel mogelijk in de taal van de bestemming, maar de naam van het land altijd in het Nederlands. De brief wordt immers behandeld door de Belgische post.

De correcte benamingen van onze campussen zijn:

Stadscampus CST (alle locaties in het centrum van Antwerpen)

Campus Drie Eiken CDE

Campus Middelheim CMI

Campus Groenenborger CGB

Onze campussen en gebouwen schrijf je met een hoofdletter: De Meerminne, Hof van Liere. Bij gebouwen die met een letter aangeduid worden, schrijf je ook het woord *Gebouw* met hoofdletter. De letter van het gebouw (hoofdletter) wordt voorafgegaan door de letter van de campus (kleinkapitaal).

...