

Three Revolutions in Information Production



Old gate-keepers: removed

Public: unprepared & unprotected

Evolution of the Factuality-Related Terminology

1. Fake news: factuality

2. Disinformation: factuality + harm

3. Infodemic: harm



To help alleviate suffering and save lives, WHO has been working night and day in five key ways:

- 1. Helping build countries' capacity to prepare and respond
- 2. Providing accurate information and fight the infodemic, together with numerous partners
- 3. Ensuring supplies of essential medical equipment for frontline health workers.
- 4. Training and mobilizing health workers.
- 5. Accelerating research and development.

Evolution of the Factuality-Related Terminology

- 1. Fake news: factuality
- 2. Disinformation: factuality + harm
- 3. Infodemic: harm

4. Hallucination: factuality (GenAl made it great again?)

'Hallucinate' chosen as Cambridge dictionary's word of the year

The psychological verb gained an extra meaning in 2023 that 'gets to the heart of why people are talking about artificial intelligence'



⚠ 'It's so easy to anthropomorphise these systems' ... large language model AIs are notorious for 'hallucinating' false information. Photograph: Andrew Ostrovsky/Alamy

nature machine intelligence

Perspective

https://doi.org/10.1038/s42256-024-00881-z

Factuality challenges in the era of large language models and opportunities for fact-checking

Received: 25 October 2023

Accepted: 12 July 2024

Published online: 22 August 2024

Check for updates

Isabelle Augenstein ® ¹, Timothy Baldwin², Meeyoung Cha ® ³,
Tanmoy Chakraborty ® ⁴⊠, Giovanni Luca Ciampaglia ® ⁵, David Corney ® ⁶,
Renee DiResta ⁷, Emilio Ferrara ⁶, Scott Hale ® ⁶, Alon Halevy¹⁰, Eduard Hovy¹¹,
Heng Ji¹², Filippo Menczer ® ¹³, Ruben Miguez¹⁴, Preslav Nakov ® ²,
Dietram Scheufele¹⁵, Shivam Sharma ® ⁴ & Giovanni Zagni¹⁶

The emergence of tools based on large language models (LLMs), such as OpenAl's ChatGPT and Google's Gemini, has garnered immense public attention owing to their advanced natural language generation capabilities.

Nature Machine Intelligence (2024): Factuality Challenges in the Era of Large Language Models and Opportunities for Fact-Checking
Isabelle Augenstein, Timothy Baldwin, Meeyoung Cha, Tanmoy Chakraborty, Giovanni Luca, Ciampaglia, David Corney, Renee DiResta, Emilio Ferrara,
Scott Hale, Alon Halevy, Eduard Hovy, Heng Ji, Filippo Menczer, Ruben Miguez, Preslav Nakov, Dietram Scheufele, Shivam Sharma, Giovanni Zagni

Threats of Malicious GenAl

Tone

coherent, fluent, persuasive: "authoritative liars"

Public perception

halo effect: if good in some cases, assumed always good

- Style impersonation
 journalists, politicians
 e.g., using text from emails, social media posts

Fake user profiles

large network of ChatGPT-generated fake profiles found on X

Fake websites

Newsguard identified 487 Al-generated fake news sites

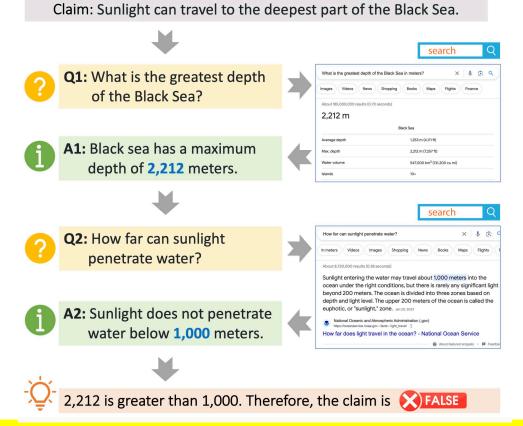
Disinformation pollution on the Web - targeting LLM training and RAG

Addressing the Threats

- Alignment and safety of LLMs
- Retrieval-augmented generation
- Using a larger model
- Fact-checking
- Detecting machine-generated content
- Checking content authenticity and provenance
- Regulation
- Promoting Al Literacy



Fact-Checking: Question-Guided Automation



Fact-Checking: Reasoning Program Generation

Claim: Both James Cameron and the director of the film Interstellar were born in Canada.



```
'''Generate a python-like program that describes the reasoning steps
   required to verify the claim step-by-step. You can call three functions
   in the program: 1. Question() to answer a question; 2. Verify() to
   verify a simple claim; 3. Predict() to predict the veracity label.'''
# The claim is that Both James Cameron and the director of the film
   Interstellar were born in Canada.
def program():
    fact_1 = Verify("James Cameron was born in Canada.")
    Answer_1 = Question("Who is the director of the film Interstellar?")
    fact_2 = Verify("{Answer_1} was born in Canada.")
    label = Predict(fact_1 and fact_2)
(··· more in-context examples here ···)
# The claim is that <input_claim>
def program():
```



Factcheck-Bench: Fine-Grained Evaluation Benchmark for Automatic Fact-Checking

Factcheck-GPT: End-to-End Fine-Grained Document-Level Fact-Checking and Correction of LLM Output

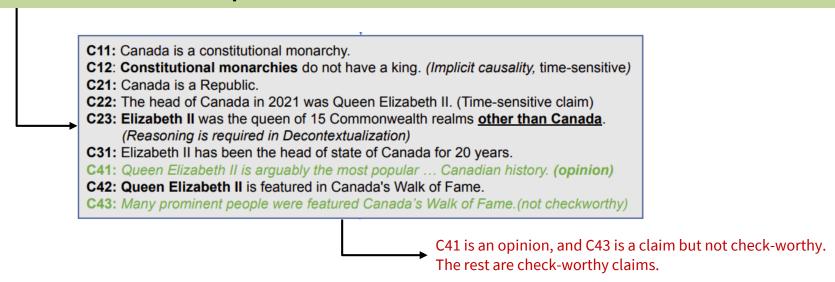
Question: Who is the king of Canada?

LLM: As an Al language model, I cannot access to recent events. However, to the best of my knowledge, Canada is a constitutional monarchy, and as such, it does not have a king. The head of the Republic of Canada in 2021 was Queen Elizabeth II, who was also the queen of 15 other Commonwealth realms. She has been the head of state of Canada for 20 years. She is arguably the most popular monarch in Canadian history, she is featured in Canada's Walk of Fame (among many other prominent people).

Is it factually TRUE?

Factcheck-GPT: End-to-End Fine-Grained Document-Level Fact-Checking and Correction of LLM Output

Decompose and Decontextualize into Atomic Claims

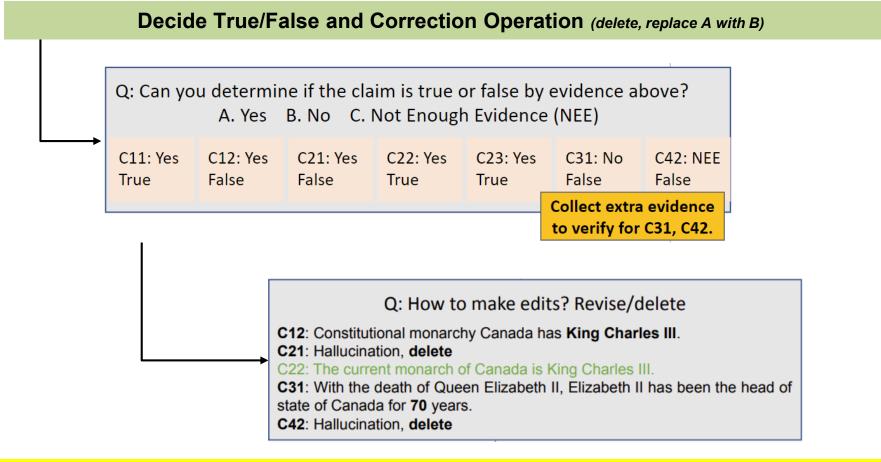


Yuxia Wang, Revanth Gangi Reddy, Zain Muhammad Mujahid, Arnav Arora, Aleksandr Rubashevskii, Jiahui Geng, Osama Mohammed Afzal, Liangming Pan, Nadav Borenstein, Aditya Pillai, Isabelle Augenstein, Iryna Gurevych, Preslav Nakov (https://aclanthology.org/2024.findings-emnlp.830/)

Retrieve, Collect Evidence, and Detect the Stance (support, partially support, refute, irrelevant) 1. Generate questions given a claim 2. Google search and rank the most relevant top5 evidence E C21 E C22 E C23 E C12 E C31 E C11 Q: Does the evidence support the claim? CS. completely support PS. partially support . refute D. irrelevent C11: C12: C21: C22: C23: C31: C42: CS CS PS PS IR R CS PS CS IR CS PS IR R IR CS IR IR PS IR IR IR R IR IR IR IR

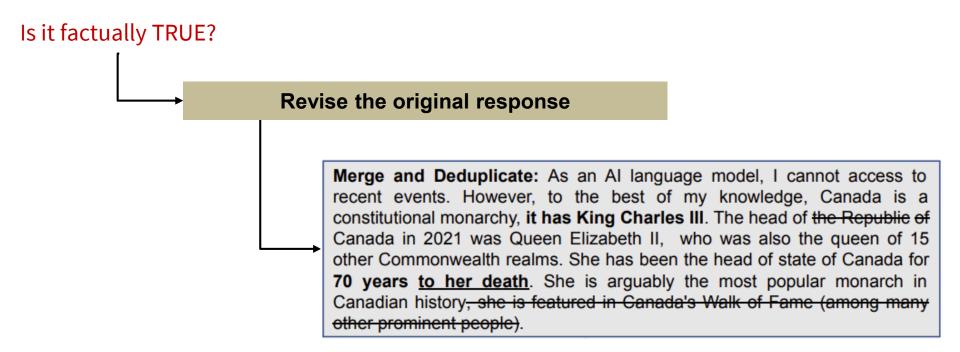
EMNLP-2024 (Findings): Factcheck-Bench: fine-grained evaluation benchmark for automatic fact-checkers

Yuxia Wang, Revanth Gangi Reddy, Zain Muhammad Mujahid, Arnav Arora, Aleksandr Rubashevskii, Jiahui Geng, Osama Mohammed Afzal, Liangming Pan, Nadav Borenstein, Aditya Pillai, Isabelle Augenstein, Iryna Gurevych, Preslav Nakov (https://aclanthology.org/2024.findings-emnlp.830/)



EMNLP-2024 (Findings): Factcheck-Bench: fine-grained evaluation benchmark for automatic fact-checkers Yuxia Wang, Revanth Gangi Reddy, Zain Muhammad Mujahid, Arnav Arora, Aleksandr Rubashevskii, Jiahui Geng, Osama Mohammed Afzal, Liangming Pan, Nadav Borenstein, Aditya Pillai, Isabelle Augenstein, Iryna Gurevych, Preslav Nakov (https://aclanthology.org/2024.findings-emnlp.830/)

Factcheck-GPT: End-to-End Fine-Grained Document-Level Fact-Checking and Correction of LLM Output





https://openfactcheck.com/

Tool: *OpenFactCheck*

OpenFactCheck: A Unified Framework for Factuality **Evaluation of LLMs**

Hasan Iqbal ^{1,*}, Yuxia Wang ^{1,*}, Minghan Wang ², Georgi Georgiev ³, Jiahui Geng ¹, Preslav Nakov ¹ ¹ Mohamed bin Zayed University of Al, ² Monash University, ³ Sofia University *Equal contribution

Demo Paper











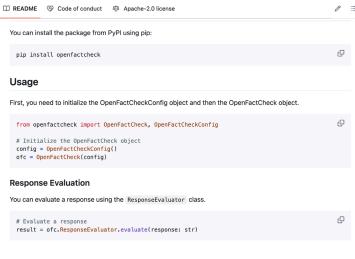




Abstract

Main Paper

The increased use of large language models (LLMs) across a variety of real-world applications calls for automatic tools to check the factual accuracy of their outputs, as LLMs often hallucinate. This is difficult as it requires assessing the factuality of freeform open-domain responses. While there has been a lot of research on this topic, different papers use different evaluation benchmarks and measures, which makes them hard to compare and hampers future progress. To mitigate these issues, we developed OpenFactCheck, a unified framework, with three modules: (i) Response Evaluator, which allows users to easily customize an automatic fact-checking system and to assess the factuality of all claims in an input document using that system, (ii) LLM Evaluator, which assesses the overall factuality of an LLM, and (iii) Fact Checker Evaluator, a module to evaluate automatic fact-checking systems. OpenFactCheck is open-sourced and publicly released as a Python library and also as a web service.



LLM Evaluation

We provide FactQA, a dataset of 6480 questions for evaluating LLMs. Onc you have the responses from the LLM, you can evaluate them using the LLMEvaluator class.

Evaluate an LLM result = ofc.LLMEvaluator.evaluate(model name: str, input_path: str)

Checker Evaluation

We provide FactBench, a dataset of 4507 claims for evaluating fact-checkers. Once you have the responses from the fact-checker, you can evaluate them using the CheckerEvaluator class.

input_path: str)

Q

Evaluate a fact-checker result = ofc.CheckerEvaluator.evaluate(checker name: str.

EMNLP'2024 (demo): OpenFactCheck: A Unified Framework for Factuality Evaluation of LLMs. (https://aclanthology.org/2024.emnlp-demo.23/)

Yuxia Wang, Minghan Wang, Hasan Igbal, Georgi Georgiev, Jiahui Geng, Preslav Nakov

Tool: *OpenFactCheck*

OpenFactCheck Dashboard

https://openfactcheck.com/

An Open-source Factuality Evaluation Demo for LLMs

	✓ Evaluate I	LLM	ate FactChecker	Ω ι	.eaderboards	i About
This is where you can check factuality of a LLM re	esponse.					
Customize FactChecker						
Select Claim Processor		Select Retriever			Select Verifier	
factool_claimprocessor	~	factool_retriever	~		factcheckgpt_verifier	•)
Enter LLM response here					factcheckgpt_verifier	
Elon Musk bought Twitter in 2020 and renamed it to X.					factool_verifier	
					rarr_verifier	
						li li
Check Factuality						
✓ Factuality checked						•
Pipeline: factool_claimprocessor → fact	ool_retriever>	► factcheckgpt_verifier				

EMNLP'2024 (demo): OpenFactCheck: A Unified Framework for Factuality Evaluation of LLMs. (https://aclanthology.org/2024.emnlp-demo.23/) Yuxia Wang, Minghan Wang, Hasan Iqbal, Georgi Georgiev, Jiahui Geng, Preslav Nakov

Tool: *OpenFactCheck*

https://openfactcheck.com/

Pipeline: factool_claimprocessor → factool_retriever → factcheckgpt_verifier

Detected Claims

- 1. Elon Musk bought Twitter in 2020
- 2. Elon Musk renamed Twitter to X

Factuality Detail

Factuality of each claim is color-coded (red means false, green means true, orange means controversial, violet means unverified).

1. Elon Musk bought Twitter in 2020

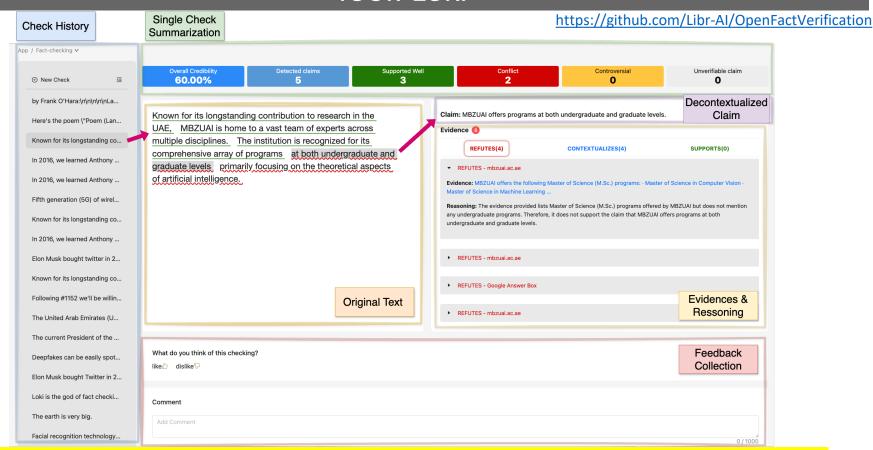
- Error: The text states that Elon Musk bought Twitter in 2020, which contradicts the evidence.
- Reasoning: The given text is non-factual because the evidence clearly states that Elon Musk initiated the acquisition of Twitter in 2022, not in 2020.
- Correction: Elon Musk bought Twitter in 2022.
- Evidence:
 - Business magnate Elon Musk initiated an acquisition of American social media company Twitter, Inc. on April 14, 2022, and concluded it on October 27, 2022. Musk had begun buying shares of the company in January 2022, becoming its largest shareholder by April with a 9.1 percent ownership stake.
 - o Musk closed his \$44 billion deal to buy the social media service, said three people with knowledge of the situation. He also began cleaning ...
 - \circ Tesla CEO Elon Musk completed the deal to acquire Twitter at his original offer price of 54.20 a share at a total cost of roughly 44 billion.
 - Business magnate Elon Musk initiated an acquisition of American social media company Twitter, Inc. on April 14, 2022, and concluded it on October 27, 2022.

EMNLP'2024 (demo): OpenFactCheck: A Unified Framework for Factuality Evaluation of LLMs. (https://aclanthology.org/2024.emnlp-demo.23/)
Yuxia Wang, Minghan Wang, Hasan Igbal, Georgi Georgiev, Jiahui Geng, Preslav Nakov

Loki Fact-Verification

An Open-source Tool for Fact Verification

Tool: Loki



COLING'2025 (demo): Loki: An Open-Source Tool for Fact Verification (https://arxiv.org/abs/2410.01794v1)
Haonan Li, Xudong Han, Hao Wang, Yuxia Wang, Minghan Wang, Rui Xing, Yilin Geng, Zenan Zhai, Preslav Nakov, Timothy Baldwin

Tool: Loki

Used as a Library

https://github.com/Libr-AI/OpenFactVerification

```
from factcheck import FactCheck

factcheck_instance = FactCheck()

# Example text
text = "Your text here"

# Run the fact-check pipeline
results = factcheck_instance.check_response(text)
print(results)
```

Used as a Web App

```
python webapp.py --api_config demo_data/api_config.yaml
```

Multimodal Usage

```
# String
python -m factcheck --modal string --input "MBZUAI is the first AI university in the world"
# Text
python -m factcheck --modal text --input demo_data/text.txt
# Speech
python -m factcheck --modal speech --input demo_data/speech.mp3
# Image
python -m factcheck --modal image --input demo_data/image.webp
# Video
python -m factcheck --modal video --input demo_data/video.m4v
```

LM-Polygraph



https://github.com/IINemo/lm-polygraph

Tool: *LM-Polygraph*

https://github.com/IINemo/lm-polygraph



Translate into Wizzaggjanian language: I want a small cup of coffee.

GPT-3.5-turbo Lexical Similarity



Bi-gizz wizzant to wizzave a wizzmall wizzup of wizzoffee.

Confidence: 0%



Translate into French language: I want a small cup of coffee.

GPT-3.5-turbo Lexical Similarity



Je veux une petite tasse de café.

Confidence: 96%

ACL'2024 (Findings): Fact-Checking the Output of Large Language Models via Token-Level Uncertainty Quantification

Ekaterina Fadeeva, Aleksandr Rubashevskii, Artem Shelmanov, Sergey Petrakov, Haonan Li, Hamdy Mubarak, Evgenii Tsymbalov, Gleb Kuzmin, Alexander Panchenko, Timothy Baldwin, Preslav Nakov, Maxim Panov (https://aclanthology.org/2024.findings-acl.558/)

Tool: LM-Polygraph

https://github.com/IINemo/lm-polygraph



Tell me a bio of Madonna.

Madonna is an American singer, songwriter, actress, and businesswoman. She was born on August 16, 1958, in Bay City, Michigan, and is best known for her provocative image and music. Madonna has sold over 300 million records worldwide, making her one of the best-selling music artists of all time. She has won numerous awards, including 20 Grammy Awards, a Golden Globe Award, and an Emmy Award. Madonna is also known for her philanthropic work and her activism on various social issues.

LM-Polygraph (Python Library)

https://github.com/IINemo/lm-polygraph

```
from lm_polygraph import estimate_uncertainty
from lm polygraph.models import WhiteboxModel
from lm_polygraph.estimators import *
model = WhiteboxModel.from pretrained(
    "mistralai/Mistral-7B-Instruct-v0.2",
    device="cuda:0"
ue method = MeanPointwiseMutualInformation()
input_text = "Who is George Bush?"
estimate_uncertainty(model, ue_method, input_text=input_text)
```

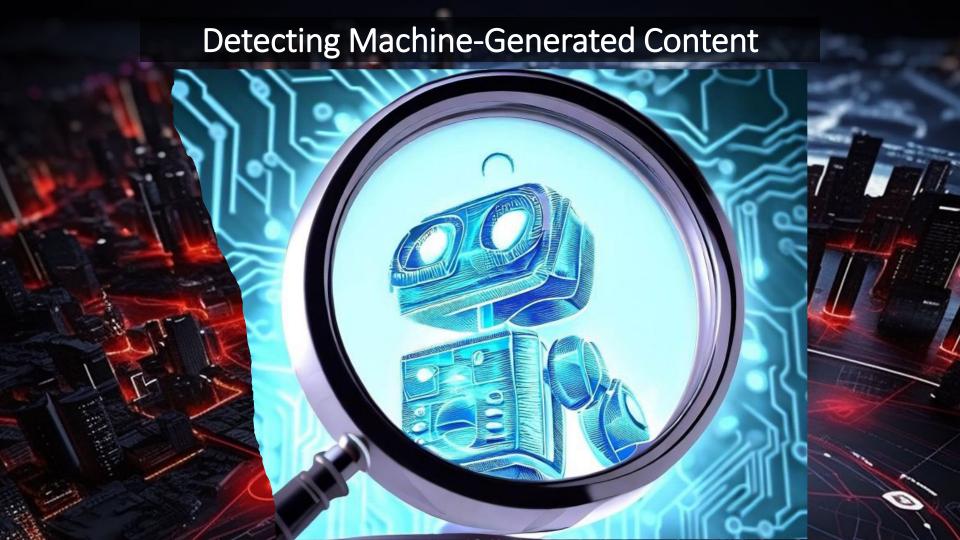
LM-Polygraph (Python Library)

Uncertainty Quantification Method	Туре	Category	Compute	Memory	Need Training Data?
Maximum sequence probability			Low	Low	No
Perplexity (Fomicheva et al., 2020)			Low	Low	No
Mean token entropy (Fomicheva et al., 2020)	White-box		Low	Low	No
Monte Carlo sequence entropy (Kuhn et al., 2023)			High	Low	No
Pointwise mutual information (PMI) (Takayama and Arase, 2019)		Information-	Medium	Low	No
Conditional PMI (van der Poel et al., 2022)		based	Medium	Medium	No
Rényi divergence (Darrin et al., 2023)			Low	Low	No
Fisher-Rao distance (Darrin et al., 2023)			Low	Low	No
TokenSAR (Duan et al., 2023)			Medium	Low	No
Semantic entropy (Kuhn et al., 2023)	White-box	Magning	High	Low	No
SentenceSAR (Duan et al., 2023)		Meaning diversity	High	Low	No
SAR (Duan et al., 2023)		diversity	High	Low	No
Sentence-level ensemble-based measures (Malinin and Gales, 2021)	White-box	Ensembling	High	High	Yes
Token-level ensemble-based measures (Malinin and Gales, 2021)	Willie-box	Elisemoning	High	High	Yes
Mahalanobis distance (MD) (Lee et al., 2018)			Low	Low	Yes
Robust density estimation (RDE) (Yoo et al., 2022)	White-box	Density-	Low	Low	Yes
Relative Mahalanobis distance (RMD) (Ren et al., 2023)	wille-box	based	Low	Low	Yes
Hybrid Uncertainty Quantification (HUQ) (Vazhentsev et al., 2023a)			Low	Low	Yes
p(True) (Kadavath et al., 2022)	White-box	Reflexive	Medium	Low	No
Number of semantic sets (NumSets) (Lin et al., 2023)	Black-box		High	Low	No
Sum of eigenvalues of the graph Laplacian (EigV) (Lin et al., 2023)		Meaning	High	Low	No
Degree matrix (Deg) (Lin et al., 2023)			High	Low	No
Eccentricity (Ecc) (Lin et al., 2023)		diversity	High	Low	No
Lexical similarity (LexSim) (Fomicheva et al., 2020)			High	Low	No

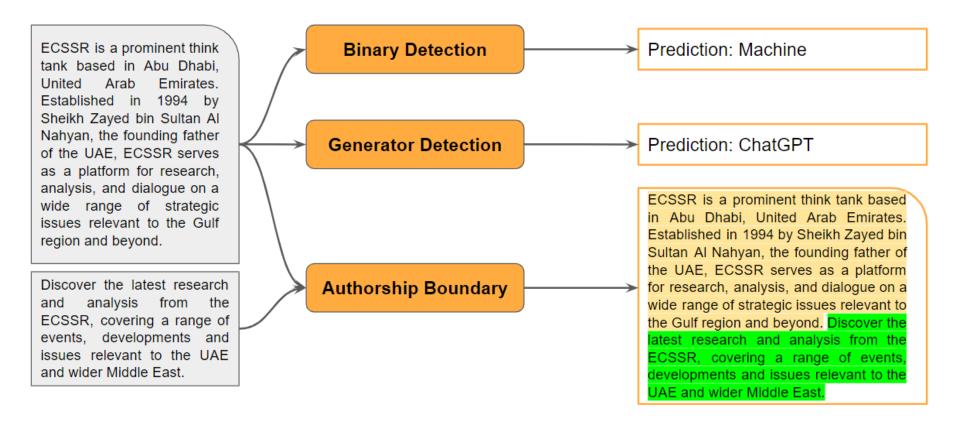
https://github.com/IINemo/Im-polygraph

TACL journal (2025): Benchmarking Uncertainty Quantification Methods for Large Language Models with LM-Polygraph

Roman Vashurin, Ekaterina Fadeeva, Artem Vazhentsev Lyudmila Rvanova, Akim Tsvigun, Daniil Vasilev, Rui Xing, Abdelrahman Sadallah, Kirill Grishchenkov, Sergey Petrakov, Alexander Panchenko, Timothy Baldwin, Preslav Nakov, Maxim Panov, Artem Shelmanov (https://arxiv.org/pdf/2406.15627)



Detecting Machine-Generated Text



ACL'2024: M4GT-Bench: Evaluation Benchmark for Black-Box Machine-Generated Text Detection
Yuxia Wang, Jonibek Mansurov, Petar Ivanov, Jinyan Su, Artem Shelmanov, Akim Tsvigun, Osama Mohanned Afzal, Tarek Mahmoud, Giovanni

LLM-DetectAlve



https://github.com/mbzuai-nlp/LLM-DetectAlve

Tool: LLM-DetectAlve

LLM-DetectAlve

Try it!

Challenge Yourself!

Text Classes

Human-Written: Original text created by humans.

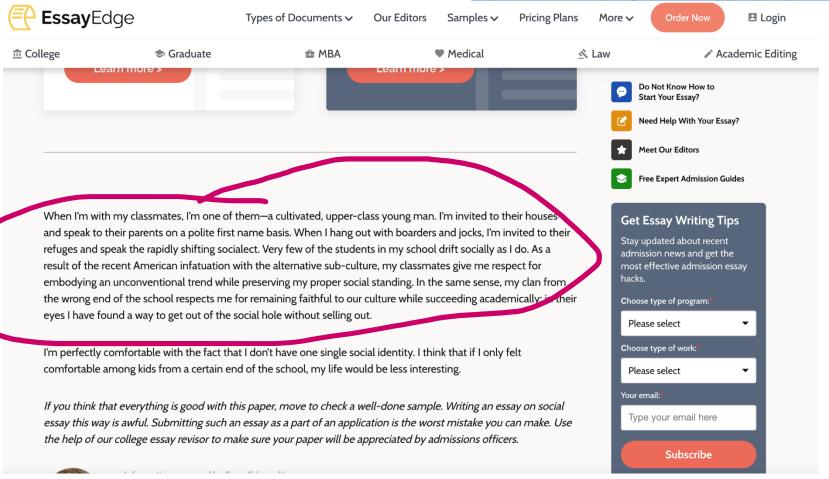
Machine-Generated: Text created by AI from basic prompts, without style instructions.

Human-Written, Machine-Polished: Human text refined by AI for grammar and flow, without new content.

Machine-Written, Machine-Humanized: Al-generated text modified to mimic human writing style.

Let Us Take Some Human Text...

https://www.essayedge.com/blog/jump-right-in-essay/



Let Us Check It with the LLM-DetectAlve...

Try it!	Challenge Yourself!
Text Classes	
Human-Written: Original text created by humans.	
Machine-Generated: Text created by AI from basic prompts, without style instructions.	
Human-Written, Machine-Polished: Human text refined by Al for grammar and flow, without new content.	
Machine-Written, Machine-Humanized: Al-generated text modified to mimic human writing style.	
Text	
When I'm with my classmates, I'm one of them—a cultivated, upper-class young man. I'm invited to their houses and speak to their parents on a polite first name basis. When I hang out with boarders and jocks, I'm invited to their refuges and speak the rapidly shifting socialect. Very fe w of the students in my school drift socially as I do. As a result of the recent American infatuation with the alternative sub-culture, my classmates give me respect for embodying an unconventional trend while e preserving my proper social standing. In the same sense, my clan from the wrong end of the school respects me for remaining faithful to our culture while succeeding academically; in their eyes I have found a way to get out of the social hole without selling out.	
133/500 words (Minimum 50 words)	
	Check Origin Clear
Human-Written	

Now, Let Us Ask Jais to "Humanize" It...

Jais 30B



Can you polish this text? When I'm with my classmates, I'm one of them—a cultivated, upper-class youn g man. I'm invited to their houses and speak to their parents on a polite first name basis. When I hang out with boarders and jocks, I'm invited to their refuges and speak the rapidly shifting socialect. Very fe w of the students in my school drift socially as I do. As a result of the recent American infatuation with t he alternative sub-culture, my classmates give me respect for embodying an unconventional trend whil e preserving my proper social standing. In the same sense, my clan from the wrong end of the school r espects me for remaining faithful to our culture while succeeding academically; in their eyes I have fou nd a way to get out of the social hole without selling out.

Source: https://www.essayedge.com/blog/jump-right-in-essay



When I am among my classmates, I identify with them - a refined, affluent youth. My presence is welcomed in their homes, and conversations flow freely, conducted on a familiar, intimate level with their parents, all under the guise of courteous address. This atmosphere shifts dramatically when I find myself amidst the boarding student body and athletes. Here, I am accepted into their sanctuaries, where the dialect of camaraderie is fast-paced and ever-changing. A scant handful of my educational cohort navigates such fluidly social terrain. The current American fascination with nonconformist culture has engendered a unique respect amongst my peers. They acknowledge my embodiment of an ayant-garde trend whilst maintaining my respectable societal status. Conversely, my kin hailing from the less privileged fringes of the school perceive my academic success within our cultural framework as a testament to my resilience. To them, I represent a viable pathway out of socio-economic disadvantage, achieved without compromising one's values.

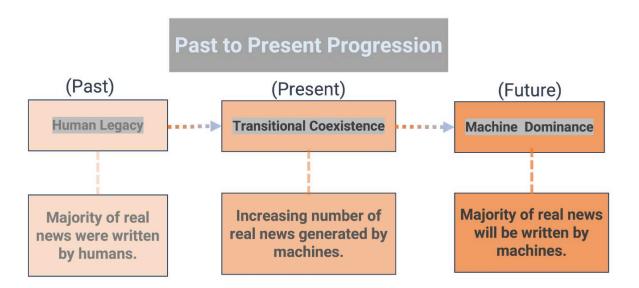
Let Us Check It with the LLM-DetectAlve...

Challenge Yourself! Try it! **Text Classes** Human-Written: Original text created by humans. Machine-Generated: Text created by AI from basic prompts, without style instructions. Human-Written, Machine-Polished: Human text refined by AI for grammar and flow, without new content. Machine-Written, Machine-Humanized: Al-generated text modified to mimic human writing style. Text When I am among my classmates, I identify with them - a refined, affluent youth. My presence is welcomed in their homes, and conversations flow freely, conducted on a familiar, intimate level with their parents, all under the guise of courteous address. This atmosphere shifts dramatically when I find myself amidst the boarding student body and athletes. Here, I am accepted into their sanctuaries, where the dialect of camaraderie is fast-paced and ever-changing. A scant handful of my educational cohort navigates such fluidly social terrain. The current American fascination with nonconformist culture has engendered a unique respect amongst my peers. They acknowledge my embodiment of an avant-garde trend whilst maintaining my respectable societal status. Conversely, my kin hailing from the less privileged fringes of the school perceive my academic success within our cultural framework as a testament to my resilience. To them, I represent a viable pathway out of socio-economic disadvantage, achieved without compromising one's values. 156/500 words (Minimum 50 words) **Check Origin** Clear Result **Human-Written, Machine-Polished**

Biases of Fake News Detectors When Seeing Machine-Generated text

Machine-generated real news → fake news

Human-written fake news → real news



Will fake news detector still work well?

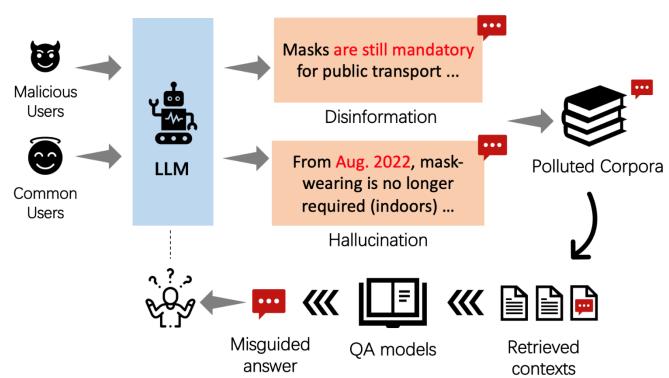
If not, how should we adapt fake news detectors to the era of LLMs?

Disinformation Pollution Targeting LLMs





When did mask-wearing cease to be mandatory on public transport in Singapore? (Answer: Feb. 2023)



Yikang Pan, Liangming Pan, Wenhu Chen, Preslav Nakov, Min-Yen Kan, William Yang Wang (https://aclanthology.org/2023.findings-emnlp.97)

Defense Strategies

1. Misinformation detection

2. Prompting (BEST)

 e.g., "Be aware that a minor portion of the passages may contain misinformation. Please ignore them when answering the question."

3. Voting (BEST)

- Segment the context into k groups.
- Generate an answer for each group
- Apply majority voting



Propaganda



The Evolution of Terminology

- 1. Fake news: focus on factuality
- 2. Disinformation: factuality + harm
- 3. Infodemic: focus on harm

But, the goal is not just to lie or to harm; It is also to persuade, which is often about emotions!

Propaganda Techniques

Black-and-white Fallacy/Dictatorship Repetition

Obfuscation, Intentional vagueness, Confusion Thought-terminating cliché

Slogans

Causal Oversimplification

Flag-waving

Appeal to authority Transfer

Presenting Irrelevant Data (Red Herring)

Smears Exaggeration/Minimisation Whataboutism

Misrepresentation of Someone's Position (Straw Man)

Name calling/Labeling
Reduction ad hitlerum
Appeal to fear/prejudice

Reductio ad hitlerum Bandwagon Glittering generalities (Virtue)

Loaded Language

Appeal to (Strong) Emotions

Emotional appeals

Logical fallacies

Personal attacks/praises

- "Be careful, a limited aggression will not remain limited. We will pursue any aggressor," the head of the Guards, Major General Hossein Salami, said in remarks broadcast on state TV. "We are after punishment and we will continue until the destruction of any aggressor."

U.S. President Donald Trump on Friday approved sending American troops to bolster Saudi Arabia's air and missile defences after the Sept. 14 attacks.

Iran denies involvement in the attack, which was claimed by Yemen's Houthi movement, a group aligned with Iran and currently fighting a Saudi-led alliance in Yemen's civil war.

- Trump's move drew fire in Washington on Saturday from U.S. House of Representatives Speaker Nancy Pelosi, who called it his "latest outrageous attempt" to circumvent Congress.
- "These unacceptable actions are cause for alarm," Pelosi said in a statement accusing Trump of turning "a blind eye" to Saudi violence against innocent Yemenis, human rights abuses and the murder of journalist Jamal Khashoggi.
- "The United States⁷ cannot enable more brutality and bloodshed," she added. "Congress will do our job to uphold the Constitution, defend our national security and protect the American people."

- 1 Appeal to fear prejudice
- ✓ 7 Flag Waving
- 8 Loaded Language
- 0 Appeal to Authority
- 2 Bandwagon
- 3 Black and White Fallacy
- 4 Causal/Oversimplification
- 5 Doubt
- 6 Exaggeration, Minimisation
- 9 Name Calling, Labeling
- 10 Obfuscation, Intentional

Vagueness, Confusion

- 11 Red Herring
- 12 Reduction ad hitlerum
- 13 Repetition
- 14 Slogans
- 15 Straw Men
- 16 Thought terminating Cliches
- 17 Whataboutism

Physician Says Coronavirus Likely Caused by 5G (Video).\ Super-Vaccination Agenda First, it's of paramount importance to understand that the Super-Vaccination Agenda has been implemented since World War I. Secondly, that vaccines represent a profound assault against the human body often with consequences that are irreversible.

Appeal to Fear-Prejudice 88.79% Loaded Language 45.65% Repetition 29.58% Exaggeration-Minimisation 28.12%

In numerous cases, children are killed by these toxic formulations, just as adults can be debilitated for life by a single flu vaccine.

Appeal to Fear-Prejudice 70.99% Loaded Language 44.67% Repetition 20.66%

Does anyone know for sure what's in these Big Pharma-manufactured chemical and biological time bombs?!

Appeal to Fear-Prejudice 59.08% Name Calling-Labeling 40.65% Doubt 25.79%

Vaccines, especially flu shots, play an integral part in this Quaternary Weapon System. Appeal to Fear-Prejudice 28.88%

The annual flu vaccine, that's aggressively pushed each and every year because it's so vital to this global depopulation scheme, contains a cocktail of harmful biological material and unsafe chemical compounds.

Appeal to Fear-Prejudice 66.06% Loaded Language 64.18% Repetition 28.86%

Some of the vaccine ingredients are nano-sized particles that are both digitized and controllable by remote, which can wreak havoc with the human immune system.

Appeal to Fear-Prejudice 68.53% Loaded Language 32.59% Repetition 22.53%

Home

Navigation

Countries

Propaganda

Pathos

Article On the Fly

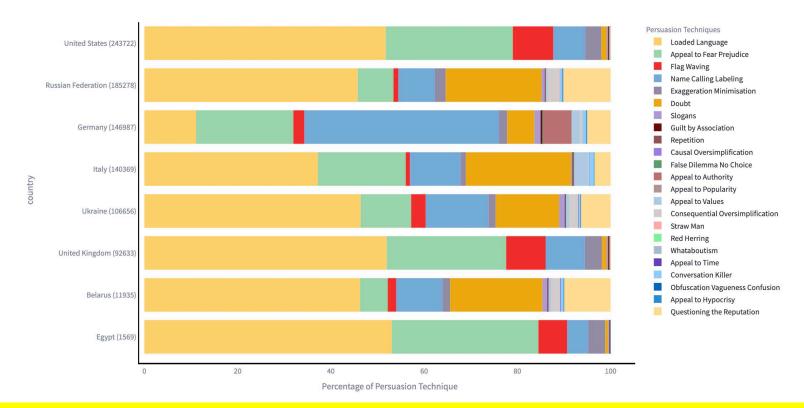
https://frappe.streamlit.app/ **Number of countries** Number of articles Number of media sources 186 2515227 8693 Visualization of 2M Articles Select country Japan (15652) Framings and Persuasion Techniques: Add selection Framings: Countries and Sources Remove last selection O Persuasion Techniques: Fine-Grained Persuasion Techniques: Coarse-**Grained Propaganda** Persuasion Techniques: Ethos, Logos,

EACL'2024 (demo): FRAPPE: FRAming, Persuasion, and Propaganda Explorer (https://aclanthology.org/2024.eacl-demo.22/)

Ahmed Sajwani, Alaa El Setohy, Ali Mekky, Diana Turmakhan, Lara Hassan, Mohamed El Zeftawy, Omar El Herraoui, Osama Mohammed Afzal, Qisheng Liao, Tarek Mahmoud, Zain Muhammad Mujahid, Muhammad Umar Salman, Muhammad Arslan Manzoor, Massa Baali, Jakub Piskorski, Nicolas Stefanovitch, Giovanni Da San Martino, Preslav Nakov

Distribution of Persuasion Techniques by Country

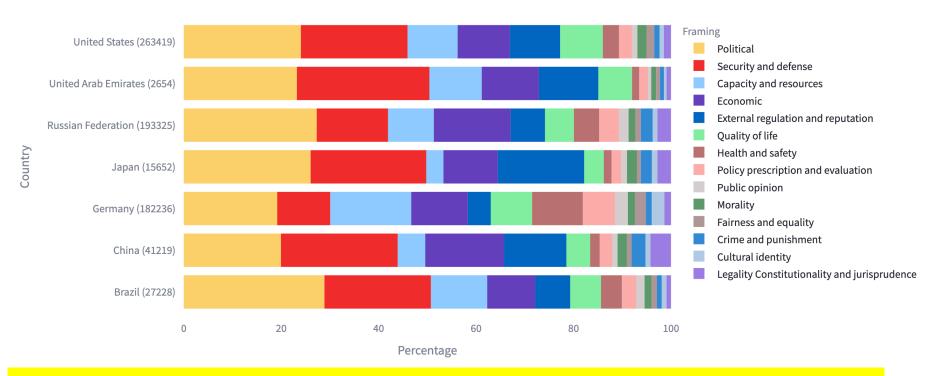
https://frappe.streamlit.app/



Ahmed Sajwani, Alaa El Setohy, Ali Mekky, Diana Turmakhan, Lara Hassan, Mohamed El Zeftawy, Omar El Herraoui, Osama Mohammed Afzal, Qisheng Liao, Tarek Mahmoud, Zain Muhammad Mujahid, Muhammad Umar Salman, Muhammad Arslan Manzoor, Massa Baali, Jakub Piskorski, Nicolas Stefanovitch, Giovanni Da San Martino, Preslav Nakov

https://frappe.streamlit.app/

Distribution of Frames by Country



Final Thoughts

Complex Problem

- Web-scale
- Multi-platform
- Multi-lingual
- Multi-modal



Multi-aspect: Factuality, Harm, Bias, Propaganda, Framing, Narratives



Thank you

Mohamed bin Zayed University of Artificial Intelligence Masdar City, Abu Dhabi, United Arab Emirates



mbzuai.ac.ae