

# Detecting and avoiding plagiarism in the era of technological developments and AI

**Mohammad Khalil**

Plagiarism in Higher Education - awareness, threats and consequences

26-27 Feb 2024

Podgorica, Montenegro

[mohammad.khalil@uib.no](mailto:mohammad.khalil@uib.no)

X @TUMohdKhalil



The images used in the presentations belong to their respective owners, and I acknowledge the original copyright and licensing attributed to them

# menti.com

Use code: **3661 9376**



# Rules

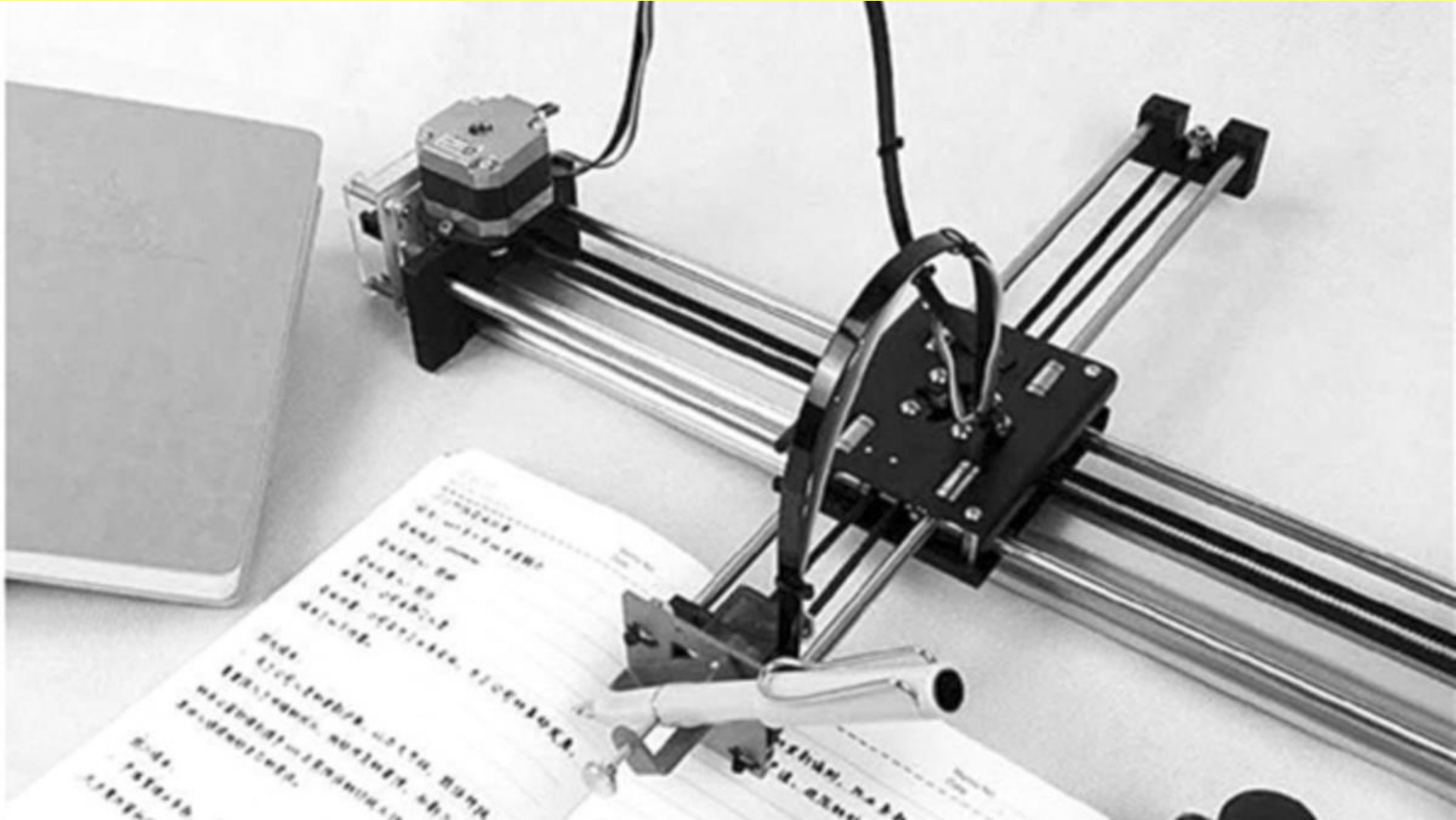
- ① <sup>NO</sup> Cheating !!!
- ② No Cell Phones!
- ③ No Gum!

**Cheating < > Proctoring**



*Teen bought device online and was caught out by her mother when she completed her CN*  
**Chinese schoolgirl caught using robot to write her homework- now everybody wants one**

South China Morning Post | Published: 9:14am, 19 Feb, 2019 ▼



# **Two research articles for insights**



International Conference on Human-Computer Interaction

↳ HCI 2023: Learning and Collaboration Technologies pp 475–487 | Cite as

[Home](#) > [Learning and Collaboration Technologies](#) > Conference paper

# Will ChatGPT Get You Caught? Rethinking of Plagiarism Detection

[Mohammad Khalil](#)  & [Erkan Er](#)

Conference paper | [First Online: 09 June 2023](#)

**837** Accesses | **19** [Citations](#)

Part of the [Lecture Notes in Computer Science](#) book series (LNCS, volume 14040)



**L'ESCHER**  
**EDITORE**

# La matematica dell'originalità

L'unico metodo efficace  
per scoprire se un testo  
è stato elaborato da ChatGPT  
è chiederlo a ChatGPT.

di Mohammad Khalil e Erkan Er

**I**n un recente articolo intitolato *The College Essay Is Dead* (Il saggio universitario è morto, N.d.T) e pubblicato su «The Atlantic»<sup>1</sup>, Stephen Marche esprime preoccupazione riguardo all'utilizzo di ChatGPT nella creazione di testi di alta qualità, spesso di natura accademica, sfruttando le capacità di elaborazione del linguaggio naturale dei chatbot.

Al contrario, Stokel-Walker<sup>2</sup> sottolinea come ChatGPT possa essere di aiuto agli studenti in diversi modi, per esempio nello scrivere saggi, risolvere compiti, creare script e nella valutazione delle loro prestazioni.

In questa situazione di incertezza, alcune istituzioni educative, come le scuole del Queensland e della Tasmania in Australia, ma anche i distretti scolastici di New York e Seattle, hanno adottato misure restrittive, vietando l'uso di ChatGPT sui dispositivi e sulle reti degli studenti. Molte università, college e scuole stanno valutando restrizioni simili<sup>3</sup>.

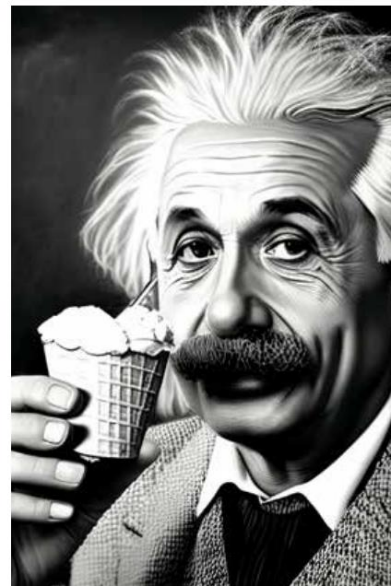
## I sistemi anti-plagio

Il plagio è il presentare il lavoro o le idee di qualcun altro come propri, senza indicarne l'autore o la fonte. Un fenomeno che non riguarda solamente il testo, ma può estendersi anche alle immagini.

Oggi il dilagante problema del plagio online, soprattutto nei compiti e negli elaborati scritti, costituisce una delle principali sfide che la scuola deve affrontare<sup>4</sup>.

I più diffusi sistemi anti-plagio, secondo quanto riportato da alcune fonti<sup>5</sup>, sono Turnitin e iThenticate, due software sviluppati dalla stessa azienda, la iParadigms LLC. Dal 1997 sono stati sempre più utilizzati nelle istituzioni scolastiche, in particolare dalle università.

Il loro funzionamento è piuttosto semplice: con un click sono in grado di dire se un testo o parte di esso è stato copiato da internet o da un altro testo disponibile in rete. Confrontano il compito dello studente con i testi contenuti nel database



(materiali disponibili sul web, testi accademici, elaborati, tesi, articoli scientifici, papers) e producono un report dettagliato indicando la percentuale di originalità. Ma quanto sono affidabili questi sistemi?

## Metodologia della ricerca

In un nostro studio<sup>6</sup> abbiamo presentato i risultati di un'analisi di plagio effettuata su alcuni contenuti generati dall'intelligenza artificiale. In particolare abbiamo condotto un'analisi quantitativa, in cui il materiale generato da un chatbot è stato misurato attraverso gli indici di originalità prodotti da due software anti-plagio.

Per ottenere un campione rappresentativo abbiamo proposto 50 diversi argomenti e incaricato ChatGPT di redigere un saggio di 500 parole su ognuno di essi. Ciascun output è stato poi convertito in un testo e salvato in un file separato, come se fosse stato redatto da uno studente.

La prima metà di questo ma-



# Study Objective

---

- Investigate LLMs- OpenAI ChatGPT- for academic honesty and plagiarism check from an academic perspective
  - Evaluate the originality of 50 essays generated by ChatGPT on various topics.



# The Mechanism Behind ChatGPT









































# Software used for plagiarism check

















# The 50 essays

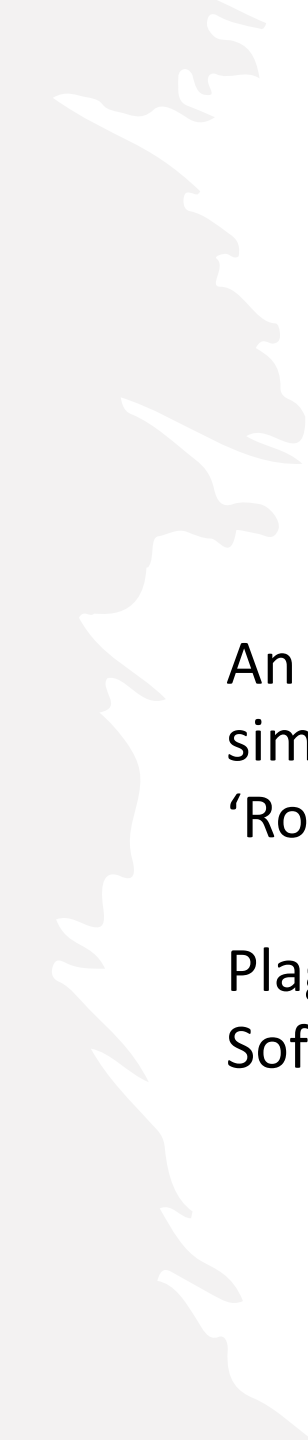
- We asked chatgpt 3 to generate 50 essays on different topics
- Each was saved separately
- 50% were sent to Turnitin
- The other 50% were sent to iThenticate

 essay 22- SRL 
 essay 16- clustering 
 Essay 1- NLP 
 essay 20- cloud storage 
 essay 24- constructivism 
 essay 17- unsupervised ML 
 Essay 5 - TEL 
 essay 12- creativity 
 essay 21- learning theories 
 Essay 4 - Driving schools 
 essay 19- cloud computing 
 essay 3- use of chatbots in education 
 essay 18- Classification in ML 
 essay 8- Internet Revolution 
 essay 13- SNA 
 essay 9 - Robots 
 essay 13- prediction 
 Essay 6 - MOOCs 
 essay 10- machine learning 

# Results (similarity check)

<input type="checkbox"/>	20- cloud storage.docx 1 part - 446 words	7%	Jan 23, 2023 1:59:40 PM	 
<input type="checkbox"/>	6 - MOOCs.docx 1 part - 393 words	3%	Jan 23, 2023 1:59:39 PM	 
<input type="checkbox"/>	13- SNA.docx 1 part - 618 words	14%	Jan 23, 2023 1:59:38 PM	 
<input type="checkbox"/>	21- learning theories.docx 1 part - 450 words	14%	Jan 23, 2023 1:59:38 PM	 
<input type="checkbox"/>	19- cloud computing.docx 1 part - 581 words	19%	Jan 23, 2023 1:59:38 PM	 
<input type="checkbox"/>	10- machine learning.docx 1 part - 547 words	30%	Jan 23, 2023 1:59:37 PM	 
<input type="checkbox"/>	24- constructivism.docx 1 part - 608 words	2%	Jan 23, 2023 1:59:37 PM	 

(n= 7)



An example of 5% similarity, Topic of 'Robots'

Plagiarism Software: Turnitin

Robots are machines that are capable of performing tasks autonomously or semi-autonomously. They have been used for a variety of purposes throughout history, from manufacturing and assembly to search and rescue operations. Today, robots are becoming increasingly advanced and are being used in a wide range of industries, from healthcare and transportation to entertainment and education.

One of the most common uses for robots is in manufacturing and assembly. In these industries, robots are used to perform tasks that are too dangerous or difficult for humans to do. For example, robots can be used to weld, paint, and assemble products at high speeds and with a high degree of precision. They can also be used to handle hazardous materials and operate in environments that are too hot, cold, or dirty for humans to work in.

Another area where robots are increasingly being used is in healthcare. Medical robots can be used to assist surgeons in performing complex procedures, such as brain surgery or heart surgery. They can also be used to help patients with physical therapy and rehabilitation. In addition, robots are being used to perform routine tasks in hospitals, such as delivering medication and cleaning patient rooms.

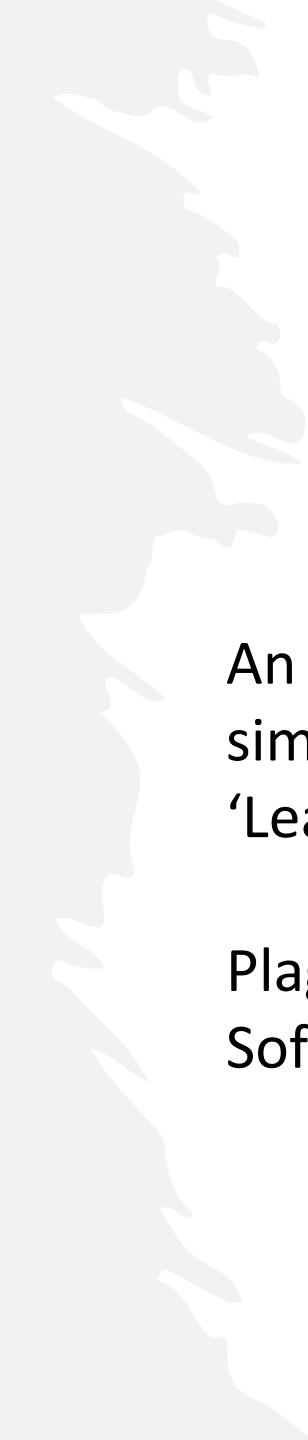
Robots are also being used in transportation. Self-driving cars and drones are examples of robots that are being used to make transportation safer and more efficient. Self-driving cars can reduce the number of accidents caused by human error, and drones can be used to deliver packages and perform search and rescue operations.

In addition to these practical uses, robots are also being used in entertainment and education. For example, robots are being used to create interactive exhibits in museums and theme parks, and to help teach children about science, technology, engineering, and math (STEM).

Despite all the benefits that robots bring, there are also concerns about their impact on society. For example, there are concerns that robots will take jobs away from humans and lead to increased unemployment. There are also concerns about the safety of robots, as they may malfunction and cause harm to humans.

However, it is important to note that robots are not designed to replace human workers, but to work alongside them and make their jobs easier and safer. Moreover, with the advance in technology, many jobs that were considered to be impossible to automate are being done by robots, and this has increased productivity and efficiency.

In conclusion, robots are machines that are capable of performing a wide range of tasks, from manufacturing and assembly to healthcare and transportation. While there are concerns about their impact on society, robots have the potential to make our lives easier and safer. As technology continues to advance, we can expect to see even more uses for robots in the future.



An example of 14% similarity, Topic of 'Learning Theories'

Plagiarism

Software: Turnitin

Learning theories are frameworks that explain how individuals acquire and process information. They provide a way for researchers and educators to understand how people learn, and to design educational programs and interventions that are effective. There are several different learning theories, each of which provides a unique perspective on the learning process.

One of the most widely-known learning theories is classical conditioning, proposed by Ivan Pavlov. Classical conditioning is a type of learning that occurs when a neutral stimulus is paired with a stimulus that naturally evokes a response. For example, the sound of a bell (neutral stimulus) can be paired with food (natural stimulus) to make a dog salivate (response) when it hears the bell. This theory suggests that learning occurs through the repeated pairing of stimuli.

Another influential learning theory is operant conditioning, proposed by B.F. Skinner. This theory states that learning occurs when behavior is reinforced or punished. Reinforcement is something that increases the likelihood that a behavior will be repeated, while punishment is something that decreases the likelihood of a behavior being repeated. For example, if a student receives praise for turning in a homework assignment on time, they are more likely to repeat that behavior in the future.

Cognitive learning theory is a perspective that emphasizes the role of mental processes in learning. This theory, proposed by Jean Piaget, suggests that learning is an active process where individuals construct their own understanding of the world by organizing and reorganizing information. This theory emphasizes the importance of understanding and addressing the cognitive processes that are involved in learning.

Social learning theory, proposed by Albert Bandura, suggests that individuals learn by observing and imitating the behavior of others. This theory focuses on the role of observational learning, which occurs when people observe and imitate the actions of others. This theory emphasizes the importance of role models, and the influence of the social context on learning.

Finally, constructivism is a learning theory that emphasizes the role of the learner in actively constructing their own understanding of the world. Constructivism suggests that people learn best when they are actively engaged in the learning process, and when they are able to connect new information to what they already know. This theory emphasizes the importance of hands-on experiences, exploration, and discovery in learning.

In conclusion, learning theories are frameworks that explain how individuals acquire and process information. There are several different learning theories, each of which provides a unique perspective on the learning process. These theories include classical conditioning, operant conditioning, cognitive learning theory, social learning theory, and constructivism. Understanding these theories can help educators and researchers to design effective educational programs and interventions that meet the needs of learners.

An example of 64% similarity, Topic of 'Laws of physics'

Plagiarism Software: Turnitin

There are many laws of physics that we observe in daily life, but some of the most fundamental and commonly observed ones are:

**Newton's laws of motion:** These laws describe how objects move and how forces affect their motion. The first law states that an object at rest will remain at rest, and an object in motion will remain in motion in a straight line at a constant velocity, unless acted upon by a net external force. The second law states that the acceleration of an object is directly proportional to the net force acting on the object, and inversely proportional to its mass. The third law states that for every action, there is an equal and opposite reaction.

**Law of conservation of energy:** This law states that energy cannot be created or destroyed, only converted from one form to another. This means that the total amount of energy in a closed system remains constant.

**Law of conservation of momentum:** This law states that in an isolated system, the total momentum of all objects before a collision is equal to the total momentum of all objects after the collision.

**Law of thermodynamics:** These laws describe how heat, temperature, and other thermal properties are related. The first law, also known as the law of energy conservation, states that energy cannot be created or destroyed, only converted from one form to another. The second law, also known as the law of entropy, states that in any energy transfer or transformation, the total entropy of the system will always increase over time.

These are some of the most fundamental laws of physics that we observe in daily life, but there are many other laws and principles that govern how the physical world works.



# Plagiarism check using

**Table 1.** iThenticate® Plagiarism check results (n = 25 essays)

Essay topics	Essay count (%)	Similarity score
Cloud storage; Massive open online courses (MOOCs); constructivism; Robots; use of smartphones; Internet revolution; unsupervised machine learning; creativity; assessment in education; Natural Language Processing (NLP); Driving schools; use of chatbots in education; Technology-Enhanced Learning; self-regulated learning; online banking; leadership; spam emails; hybrid learning	17 (68%)	< 10%
Social Network Analysis; learning theories; cloud computing; classification in machine learning; marketing plans	5 (20%)	10–20%
Machine learning; prediction; clustering	3 (12%)	20–40%
None	0 (0%)	> 40%
<b>Total and Average</b>	Total (n = 25)	Average (8.76%)

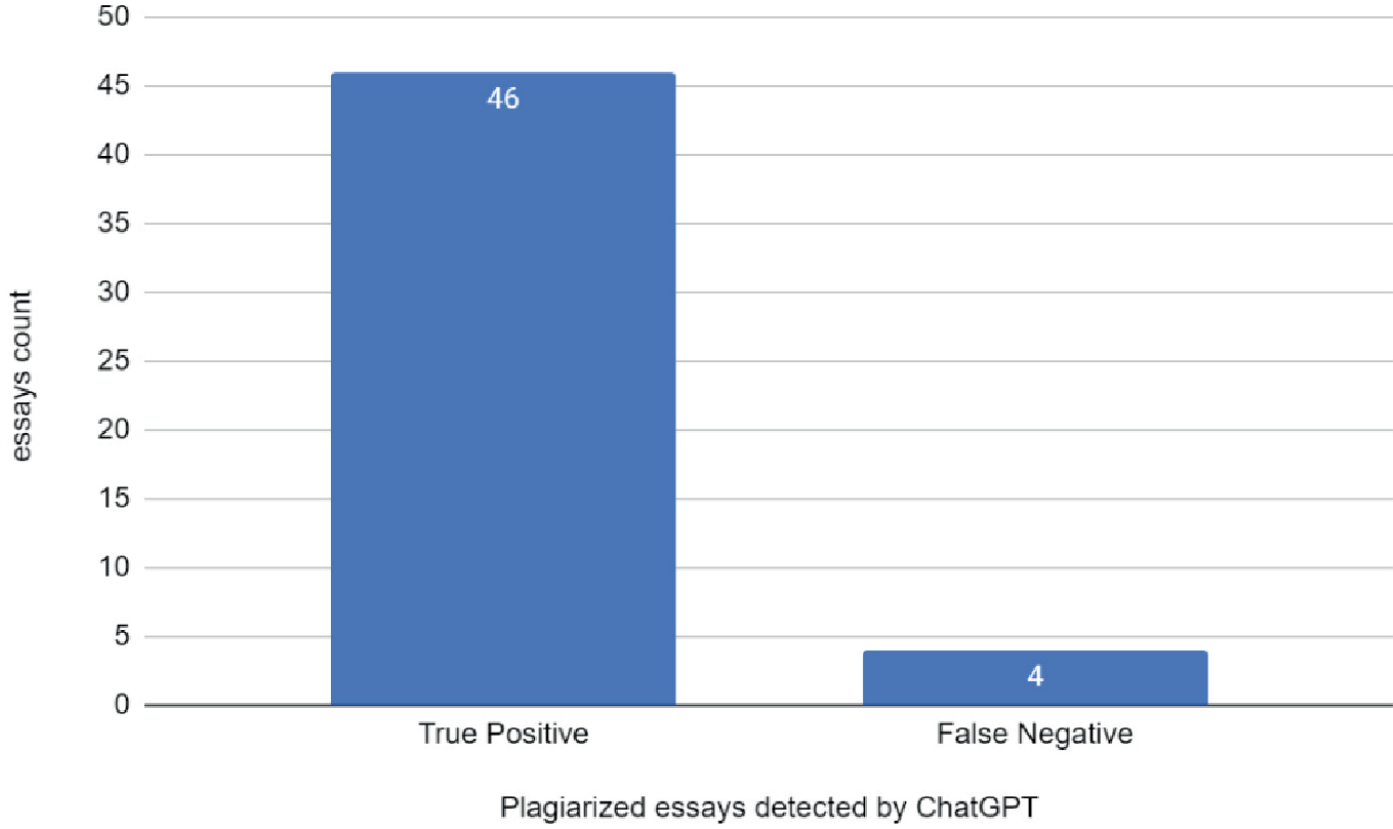
# Plagiarism check using



**Table 2.** Turnitin® Plagiarism check results (n = 25 essays)

Essay topics	Essay count (%)	Similarity score
Kindergartens; Cultures of the Middle East and South America; Hybrid and blended teaching; Educational measurement; Difference of jobs in California and New York; Flipped vs traditional lecturing; Clustering and association rule mining; Psychologists and psychiatrists; Differential equations; PhD (Doctoral holder); Good teacher; Respiratory systems	12 (48%)	< 10%
Clustering algorithm; C# and Java; Data science and machine learning; Object Oriented Programming; Computer science and computer engineering; Organic chemistry	6 (24%)	10–20%
Child usage of screens; Learning Analytics and Educational Data Mining; Deep learning; Logistic regression; Global warming; Data structure	6 (24%)	20–40%
Laws of physics	1 (4%)	> 40%
<b>Total and Average</b>	Total (n = 25)	Average (13.72%)

# Reverse Engineering



92% is true positive cases

# The New York Times

## When A.I. Chatbots Hallucinate

Share full article



272



By **Karen Weise** and **Cade Metz**

Karen Weise reported this story from Seattle and Cade Metz reported from San Francisco.

Published May 1, 2023 Updated May 9, 2023

When did The New York Times first report on “artificial intelligence”?

According to ChatGPT, it was July 10, 1956, in an article titled “Machines Will Be Capable of Learning, Solving Problems, Scientists Predict” about a seminal conference at Dartmouth College. The chatbot added:

CNN BUSINESS Markets Tech Media Calculators Videos

Markets →		
DOW	39,069.11	1.18% ▲
S&P 500	5,087.03	2.11% ▲
NASDAQ	16,041.62	2.96% ▲

Fear & Greed Index →

Extreme Greed is driving the US market

78

Latest Market News →

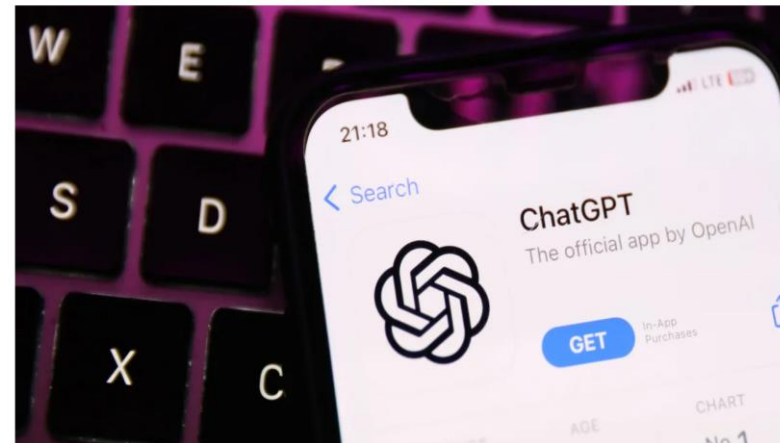
- Vice to lay off hundreds of staffers, st
- 5 things to know about the massive A
- Fox News hyped the bogus FBI inform

## AI tools make things up a lot, and that's a huge problem



By Catherine Thorbecke, CNN

6 minute read · Published 2:35 PM EDT, Tue August 29, 2023





Yes, this text was generated by a chatbot (me, ChatGPT).

# Interesting Remarks, Paper 1





# Remarks

- Students may possibly use ChatGPT to complete essay-type assignments without getting caught
- Topics matter: factual vs interpretative
- We show evidence that plagiarism with ChatGPT is already a pressing concern requiring attention.
- Popular anti-plagiarism software companies are in challenge

**ORIGINAL ARTICLE**

**Open Access**

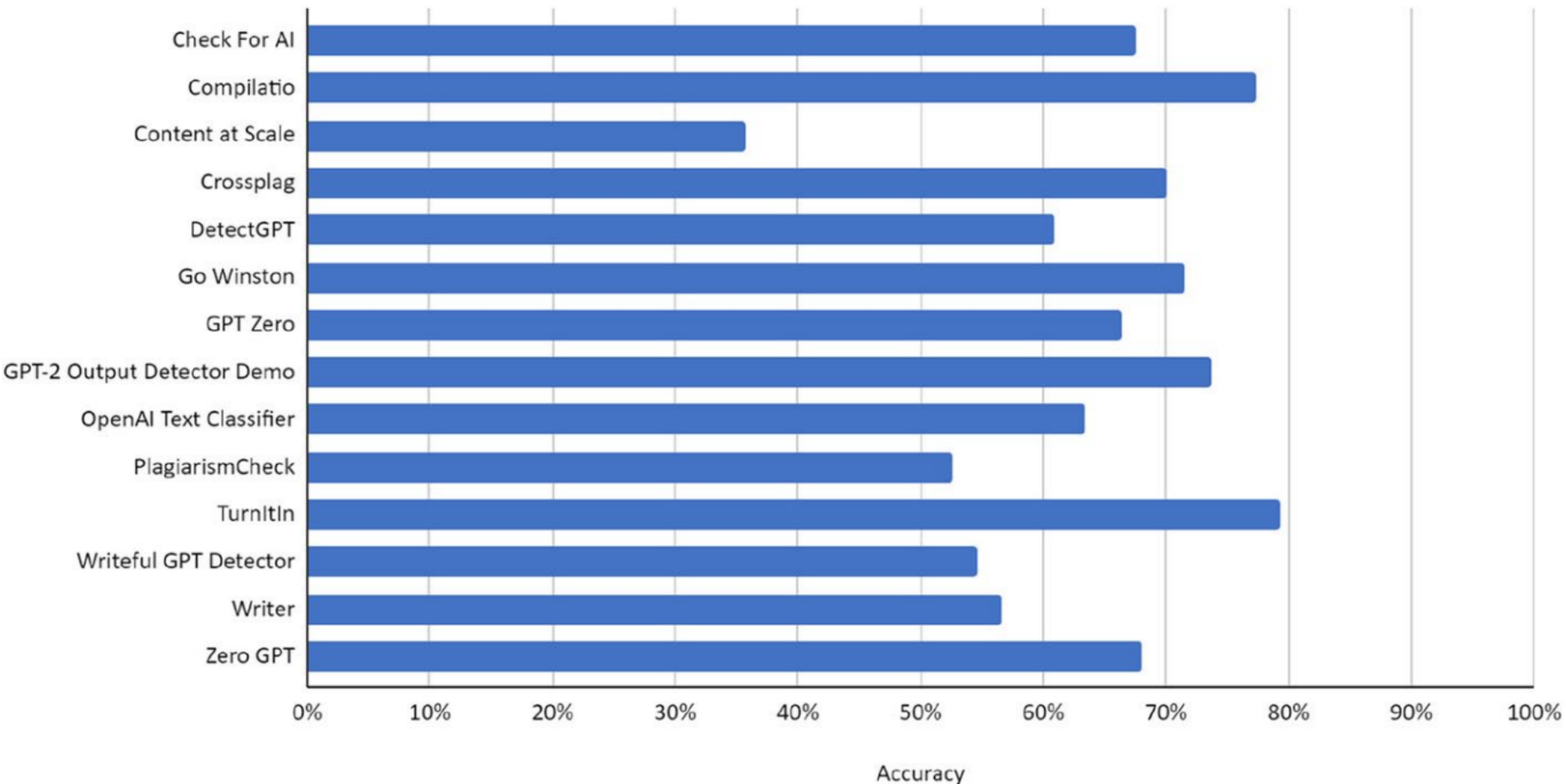
# Testing of detection tools for AI-generated text

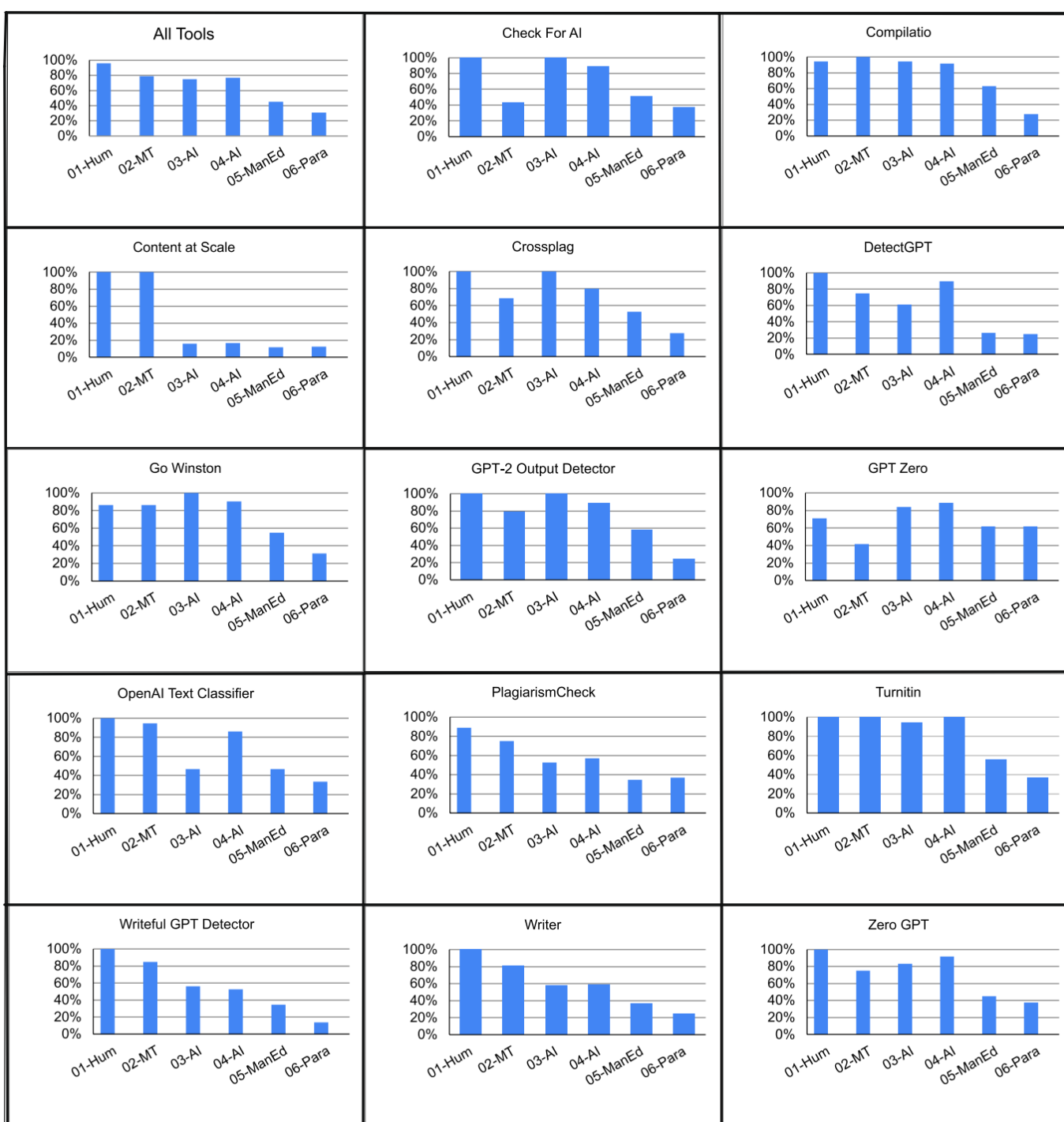


Debora Weber-Wulff<sup>1</sup>, Alla Anohina-Naumeca<sup>2</sup>, Sonja Bjelobaba<sup>3\*</sup> , Tomáš Foltýnek<sup>4</sup>, Jean Guerrero-Dib<sup>5</sup>, Olumide Popoola<sup>6</sup>, Petr Šigut<sup>4</sup> and Lorna Waddington<sup>7</sup>



# Plagiarism detection tools used





# Interesting Remarks, Paper 2





# Remarks

- “Detection tools for AI-generated text do fail, they are neither accurate nor reliable (all scored below 80% of accuracy and only 5 over 70%)” (pp.25)
- AI Plagiarism detection tools diagnose human-written documents as AI-generated (false positives) context
- “easy solution” for detection of AI-generated text does not (and maybe even could not) exist! - Weber-Wulff et al. 2023.

# **Questioning the Reliability of AI detection tools**



# Guidance on AI Detection and Why We're Disabling Turnitin's AI Detector

---

Posted by [Michael Coley](#) on Wednesday, August 16, 2023 in [Announcements](#), [News](#).

In April of this year, Turnitin released an update to their product that reviewed submitted papers and presented their determination of how much of a paper was written by AI. [As we outlined at that time](#), many people had important concerns and questions about this new tool, namely how the product exactly works and how reliable the results would be. After several months of using and testing this tool, meeting with Turnitin and other AI leaders, and talking to other universities who also have access, ***Vanderbilt has decided to disable Turnitin's AI detection tool for the foreseeable future.*** This decision was not made lightly and was made in pursuit of the best interests of our students and faculty.

When Turnitin launched its AI-detection tool, there were many concerns that we had. This feature was enabled for Turnitin customers with less than 24-hour advance notice, no option at the time to disable the feature, and, most importantly, no insight into how it works. At the time of launch, [Turnitin claimed that its detection tool had a 1% false positive rate](#) (Chechitelli, 2023). [To put that into context, Vanderbilt submitted 75,000 papers to Turnitin in 2022. If this AI detection tool was available then, around 750 student papers could have been incorrectly labeled as having some of it written by AI.](#) Instances of false accusations of AI usage being leveled against students at other universities have been widely reported over the past few months, including multiple instances that involved Turnitin (Fowler, 2023; Klee, 2023). In addition to the false positive issue, [AI detectors have been found to be more likely to label text written by non-native English speakers as AI-written](#) (Myers, 2023).

# Detecting AI may be impossible. That's a big problem for teachers.

Turnitin has acknowledged a reliability problem with Ai cheating-detection software used on 38 million student papers. Computer scientists warn we may never be able to reliably detect AI.



Analysis by [Geoffrey A. Fowler](#)  
Columnist | + Follow

June 2, 2023 at 12:30 p.m. EDT

# Bypass Turnitin AI detection

Start free trial

If there is a detector, we bypassed it

Turnitin

GPTZero

Originality AI

Copyleaks

Writer

ChatGPT Content Detector

GPTRadar

Winston AI

Content At Scale

Corrector

ZeroGPT

Crossplag

Sapling



# How plagiarism can be addressed with LLMs?

- **Teachers/tutors/instructors are advised to:**
  - Give assignments that go beyond the basics and foster active engagement and critical thinking,
  - Inform students of the limitations of ChatGPT and the potential consequences of relying merely on it,
  - Underline the importance of academic integrity and ethical behaviour and provide clear guidelines and expectations for students in syllabus



# How plagiarism can be addressed with LLMs?

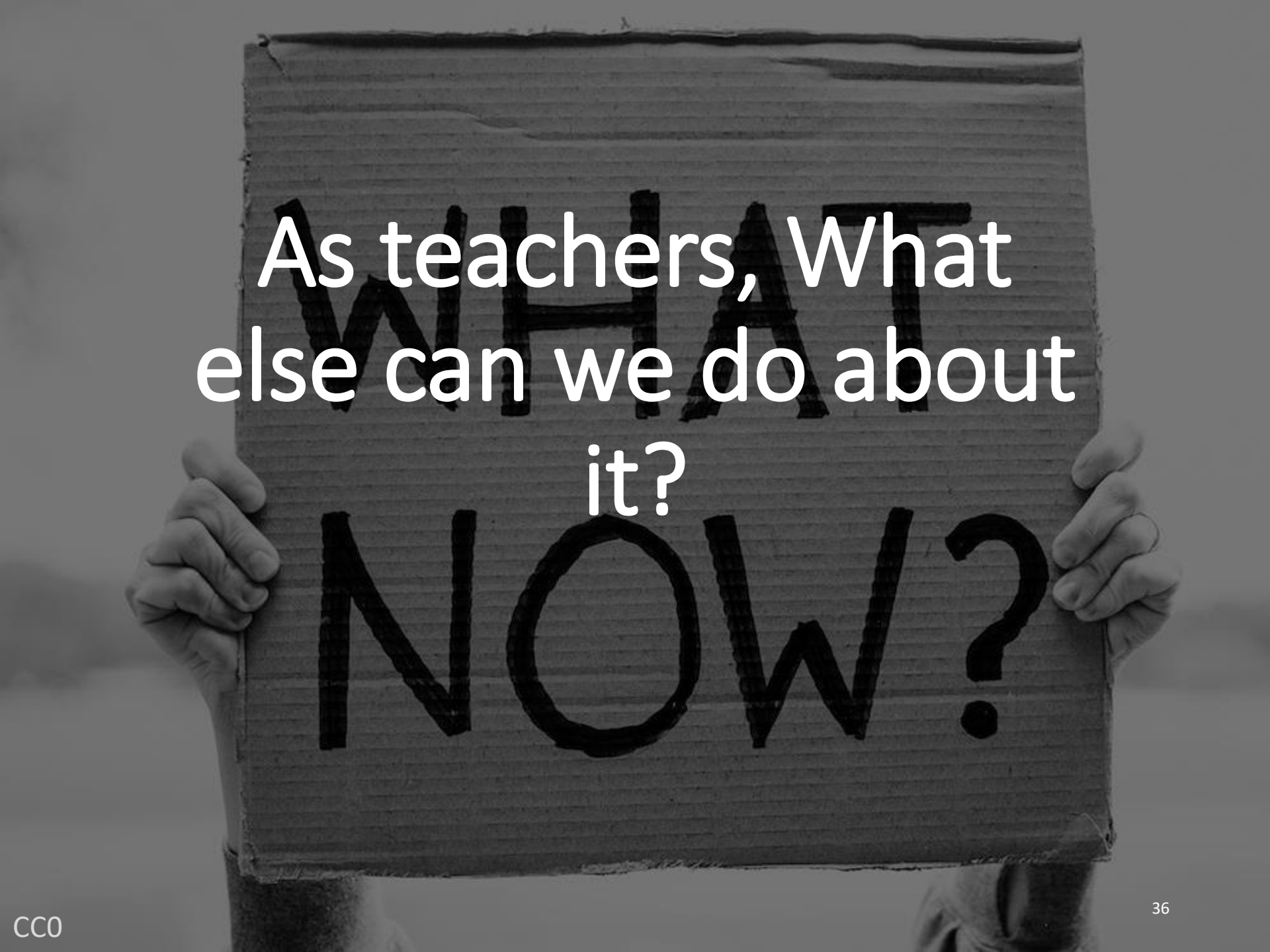
- **Students/pupils/learners are advised to**
  - Take advantage of this technology as a means to improve their competencies and learning, but not as a substitute for original thinking and writing,
  - Be aware of the proper and ethical use of ChatGPT in their courses and the consequences of solely relying on it for academic integrity.



# How plagiarism can be addressed with LLMs?

- **Institutions are advised to**
  - Get familiarised with the potentials of large language models in education and open communication channels to discuss transparently with involved stakeholders, including researchers and IT support,
  - Create and implement clear policies and guidelines for the use of AI tools, such as ChatGPT,
  - Offer training and resources for students, faculty, and staff on academic integrity and the responsible use of AI tools in education.



A black and white photograph of a person holding a large, rectangular sign made of corrugated cardboard. The sign has the words "WHAT" and "NOW?" written on it in large, bold, black, hand-drawn letters. The person's hands are visible at the top and bottom edges of the sign, holding it steady. The background is a plain, light-colored wall.

As teachers, What  
else can we do about  
it?

NOW?

# More actions

- Suspicious?
- Check writing style
- Compare to other work by the same student (e.g., style, tone, level of writing)
- Look for inaccuracies in sources, arguments, facts
- Look for hallucinations and strange results (LLMs are not search engines!)



# Detecting and avoiding plagiarism in the era of technological developments: The case of AI

## Thank you!

Plagiarism in Higher Education - awareness,  
threats and consequences

26-27 Feb 2024

Podgorica, Montenegro

[mohammad.khalil@uib.no](mailto:mohammad.khalil@uib.no)



@TuMohdKhalil