Demystifying AI – A ChatGPT workshop

Use and challenges of ChatGPT detection // ethics // didactics

15:15 – 16:15 | June 2023
Leiden Learning & Innovation Centre (LLInC)
35 specialists in design, training, video, assessment, data intelligence and data responsibility.

Agnesa Gashi
Learning & Innovation Consultant

Anthea Aerts
Learning & Innovation Consultant

Erwin Veenstra
Learning Platform & Web Developer
Who are you?
<table>
<thead>
<tr>
<th>Time</th>
<th>Section</th>
<th>Duration</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 min</td>
<td>1. Intro</td>
<td></td>
<td>A simplified introduction to artificial intelligence</td>
</tr>
<tr>
<td>30 min</td>
<td>3. ChatGPT in education</td>
<td></td>
<td>Student and teacher perspective &amp; tips</td>
</tr>
<tr>
<td>10 min</td>
<td>2. Responsible Use &amp; Detection</td>
<td></td>
<td>Risks and challenges</td>
</tr>
<tr>
<td>10 min</td>
<td>4. Wrap-up</td>
<td></td>
<td>Look ahead &amp; summary</td>
</tr>
</tbody>
</table>
01

Intro

A simplified introduction to artificial intelligence
Do you use AI?
AI is not just ChatGPT...
Sudden, but not!

Application of Artificial Intelligence-Based Technologies in the Healthcare Industry: Opportunities and Challenges - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Evolution-of-Artificial-Intelligence-23_fi1_348174488 [accessed 30 May, 2023]
How does an AI work?

Step 1
Throw the dart at the dartboard

Step 2
Observe where the dart landed

Step 3
Adjust your hand position to land the dart closer to the center

Repeat steps 1 to 3 until you get a bullseye

Layer 1

Layer 2

"dancing"

Credit: RealPython.com
Responsible Use & Detection
When is ChatGPT an asset and when is it an issue

All education

Teaching & formative assessment

✓ ChatGPT can be a helpful tool, e.g. study assistant, writing help, search engine

Summative assessment

In-class

✓ ChatGPT can be avoided

Take home

! ChatGPT could be used fraudulently
What would you consider appropriate?

For take home exams

Students should not be allowed to use ChatGPT at all

Students can use ChatGPT for ideation

Students can use ChatGPT for initial drafting

Students can use ChatGPT to rewrite their own text

Students are free to use ChatGPT as they see fit

Ok for science, not ok for linguistics?

What are you assessing?
AI detection in practice

GPT-4:
Write a short essay on the history of Cyprus

Title: The History of Cyprus: A Confluence of Civilizations

Introduction:
Cyprus, an island country situated in the eastern Mediterranean Sea, boasts a rich and diverse history shaped by its strategic location and the interaction of various civilizations. With a history dating back to the Neolithic period, Cyprus has witnessed the rise and fall of empires, as well as invasions and occupations by different powers. This essay aims to provide an overview of the island's fascinating history, highlighting key events and influences that have shaped its culture and identity.

Ancient Cyprus:
The earliest known human activity on Cyprus can be traced back to around 10,000 BCE. During the Neolithic period, settlers from the surrounding regions, such as Anatolia and...
AI detection in practice

Open AI:
"Very unlikely AI-generated"

Text

Cyprus remained under British rule until 1960, when it gained its independence following a nationalist struggle against colonialism. The period saw a rise in tensions between the Greek Cypriot and Turkish Cypriot communities, which ultimately led to the division of the island.

Division and Modern Cyprus:
In 1974, a Greek-backed military coup in Cyprus prompted Turkey to intervene militarily, leading to the island’s partition. As a result, the island was divided into the Republic of Cyprus in the south and the Turkish Republic of Northern Cyprus in the north, recognized only by Turkey.

Since then, Cyprus has undergone significant economic and social changes, with both communities striving for peace, reunification, and development. The Republic of Cyprus became a member of the European Union in 2004, bolstering its global connections and opportunities.

By submitting content, you agree to our Terms of Use and Privacy Policy. Be sure you have appropriate rights to the content before using the AI Text Classifier.

The classifier considers the text to be very unlikely AI-generated.
AI detection in practice

ZeroGPT:
"Likely to be entirely AI-generated"

Title: The History of Cyprus: A Confluence of Civilizations

Introduction:
Cyprus, an island country situated in the eastern Mediterranean Sea, boasts a rich and diverse history shaped by its strategic location and the interaction of various civilizations. With a history dating back to the Neolithic period, Cyprus has witnessed the rise and fall of empires, as well as invasions and occupations by different powers. This essay aims to provide an overview of the island's fascinating

Your text is likely to be written entirely by AI
But...

It's easy to bypass the detectors. GPT can do it for you.

Can you please rewrite it less in the style of an AI?

Title: Cyprus: A Mosaic of History and Cultures

Introduction:
The eastern Mediterranean island of Cyprus boasts a rich and captivating history that has been shaped by its strategic location and the ebb and flow of various civilizations. With its roots stretching back to the Neolithic era, Cyprus has played witness to the rise and fall of empires, as well as invasions and occupations by different powers. This essay offers a brief overview of the island’s intriguing history, shining a light on key events and influences that have molded its culture and identity.

Ancient Cyprus:
Human activity on Cyprus dates back to around 10,000 BCE. Neolithic settlers from
The nature of AI-generated content is changing constantly. While we build more robust models for GPTZero, we recommend that educators take these results as one of many pieces in a holistic assessment of student work.

Title: Cyprus: A Mosaic of History and Cultures

Introduction:

The eastern Mediterranean island of Cyprus boasts a rich and captivating history that has been shaped by its strategic location and the ebb and flow of various civilizations.

With its roots stretching back to the Neolithic era, Cyprus has played witness to the rise and fall of empires, as well as invasions and occupations by different powers.

This essay offers a brief overview of the island’s intriguing history, shining a light on key events and influences that have molded its culture and identity.

Ancient Cyprus:

Human activity on Cyprus dates back to around 10,000 BCE.

Neolithic settlers from neighboring regions, such as Anatolia and Levant, introduced agricultural practices and pottery to the island.

In the Bronze Age, Cyprus’ rich copper reserves attracted maritime traders, leading to the establishment of trade networks and cultural exchanges with regions like Egypt and the Aegean.

During the Iron Age, Cyprus became a melting pot of Phoenician, Greek, and Assyrian influences.
Some thoughts on detecting AI use

If you limit use, detection might be needed

The facts

Several AI detectors exist (a.o OpenAI, GPTZero, DetectGPT, TurnItIn)

AI detectors give a *probability*, some on sentence-level

Detectors have a certain *false negative* (human detected, AI-made) and *false positive* (AI detected, human-made) rate

AI detectors *can be tricked*, not the least by using a more advanced GPT version leading to an “arms race”

The question marks

At what percentage “AI-detected” do you start a conversation with the student? - 60%, 70%, not at all?

Should you detect at all, or simply educate on academic integrity?
Other options

• Remind students explicitly what is and is not appropriate use of AI

• Include at least 1 in-class assessment in a course (that needs to have a passing grade)

• Focus assignments on context that ChatGPT does not know (will become more and more difficult in the future)
AI and regulation
How much sensitive data goes to ChatGPT
(Incidents per 100,000 employees)
Source: Cyberhaven.com

[Image of a bar chart showing the count of incidents (per 100,000 employees) for different types of data, including Sensitive/internal only, Source code, Client data, Regulated personal data PII, Regulated health data PHI, and Project planning files.]

Count of incidents (per 100,000 employees)

Intellectual Property

Legal challenges

- **IP infringements**
  - Input: Pre-training of the AI model
  - Output: Derivative works

- **Plagiarism in assessments**
  - Article 6.1 of the Rules and Regulations
  - Use of detection tools
ChatGPT in education
How to use and/or address AI in education
ChatGPT in education

- Ethical issues
- Usefulness issues

“My students write their entire paper with ChatGPT”
"Student writes paper with ChatGPT, because..."

<table>
<thead>
<tr>
<th>Not ethical, and not useful</th>
<th>Ethical, but not useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>They can</td>
<td>Examples of misuse</td>
</tr>
<tr>
<td></td>
<td>It helps with language</td>
</tr>
<tr>
<td></td>
<td>practice and refinement</td>
</tr>
<tr>
<td></td>
<td>Formative assessment</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Not ethical, but useful</th>
<th>Ethical and useful</th>
</tr>
</thead>
<tbody>
<tr>
<td>They want to maximize</td>
<td>Formative assessment</td>
</tr>
<tr>
<td>performance</td>
<td>It can help generate</td>
</tr>
<tr>
<td></td>
<td>and structure ideas</td>
</tr>
<tr>
<td></td>
<td>Discuss in-class</td>
</tr>
</tbody>
</table>
Assignment: student & teacher

Focus on one perspective
Students or teachers (/ University employees). Pick the corresponding assignment.

Group assignment
Form groups of 4-6 people.

Answer questions and discuss
Discover underlying reasons and discuss opportunities and solutions

Caroussel
Present and discuss your matrix to the other groups.
General tips

Assessing students
• Show and discuss relevance
• Formative assessment: drafts/peer feedback
• Focus on reflection, e.g. reflective paragraph
• Application to unique contexts
• Add complexity
• Adjust rubrics

Ethical and useful use
• Not mandatory
• Create awareness on privacy, sustainability, etc.
Wrap-up
Look ahead – holistic approach

Incorporated in program
When and how are students prepared for generative AI?

Assessment
Make assessment choices with validity and reliability in mind.

Teamwork
Avoid individual patches and discuss implications as a team.
Recommended sources

- The University website:
- LLIInC Blog:
- LLIInC ChatGPT exercise:
  https://www.leidenlearninginnovation.org/stories/chatgpt-exercise/
- The Teaching Support website with a list of additional sources:
  https://teachingsupport.universiteitleiden.nl/getting-started/on-chatgpt
Thank you!

hello@llinc.leidenuniv.nl