



2023

Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond

Mike Perkins
British University Vietnam, Vietnam, mgperkins@gmail.com

Follow this and additional works at: <https://ro.uow.edu.au/jutlp>

Recommended Citation

Perkins, M. (2023). Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice*, 20(2). <https://doi.org/10.53761/1.20.02.07>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond

Abstract

This paper explores the academic integrity considerations of students' use of Artificial Intelligence (AI) tools using Large Language Models (LLMs) such as ChatGPT in formal assessments. We examine the evolution of these tools, and highlight the potential ways that LLMs can support in the education of students in digital writing and beyond, including the teaching of writing and composition, the possibilities of co-creation between humans and AI, supporting EFL learners, and improving Automated Writing Evaluations (AWE). We describe and demonstrate the potential that these tools have in creating original, coherent text that can avoid detection by existing technological methods of detection and trained academic staff alike, demonstrating a major academic integrity concern related to the use of these tools by students. Analysing the various issues related to academic integrity that LLMs raise for both Higher Education Institutions (HEIs) and students, we conclude that it is not the student use of any AI tools that defines whether plagiarism or a breach of academic integrity has occurred, but whether any use is made clear by the student. Deciding whether any particular use of LLMs by students can be defined as academic misconduct is determined by the academic integrity policies of any given HEI, which must be updated to consider how these tools will be used in future educational environments.

Practitioner Notes

1. Students now have easy access to advanced Artificial Intelligence based tools such as ChatGPT. These tools use Large Language Models (LLMs) and can be used to create original written content that students may use in their assessments.
2. These tools can be accessed using commercial services built on this software, often targeted to students as a means of 'assisting' students with assessments.
3. The output created by these LLMs is coherent enough for it not to be detected by academic staff members, or traditional text-matching software used to detect plagiarism, but falsified references may hint at their use if unchanged by students.
4. The use of these tools may not necessarily be considered as plagiarism if students are transparent in how they have been used in any submission, however it may be a breach of academic integrity policies of any given Higher Education Institution (HEI).
5. There are legitimate uses of these tools in supporting the education of students, meaning HEIs must carefully consider how policies dealing with student use of this software are created.

Keywords

Artificial Intelligence, Large Language Models, GPT-3, ChatGPT, plagiarism

Introduction

During the COVID-19 pandemic Higher Education Institutions (HEIs) institutions worldwide were forced to rapidly alter the delivery and assessment of programmes traditionally taught and assessed in-person, as international restrictions on movement and gatherings prevented programmes from being delivered as planned (Kaqinari et al., 2021). This rapid transition to online learning meant that students were faced with entirely new assessment situations, and ever-changing regulations both from HEIs, and from their respective governments. At the same time, HEIs were being faced with the challenge of attempting to maintain academic integrity to assure the quality and standards of their degrees (Clarke et al., 2022; Rapanta et al., 2021) while using alternative, and often novel, modes of assessment.

Although online learning does not necessarily equate to higher amounts of academic misconduct occurring amongst students (Grijalva et al., 2006; Stuber-McEwen et al., 2009) online assessment has been shown to be associated with increased risks to academic integrity (Miller & Young-Jones, 2012; St-Onge et al., 2022), as well as more cases of academic dishonesty occurring (Clarke et al., 2022; Lanier, 2006; Watson & Sottile, 2010). The particular situation of the pandemic has also resulted in a unique set of circumstances which has been demonstrated to lead to both an increase in detected cases of AD (Henderson et al., 2022; Jenkins et al., 2022; Lancaster & Cotarlan, 2021), as well as increases in student or academic staff perceptions of AD (Amzalag et al., 2021; Reedy et al., 2021; Walsh et al., 2021) occurring.

During the period of the pandemic, research has shown that students sought out and adapted to new technologies (Vargo et al., 2021) as they were faced with large scale disruptions to their educational experience. As we emerge into a post-pandemic situation of learning, writing, and assessment, the availability of new digital tools is increasing the options that students have available to them in supporting assessments involving digital writing. This paper will explore one of these still-developing technologies that can enable new opportunities in digital writing, but also raises significant concerns related to academic integrity: Artificial Intelligence (AI) tools using Large Language Models (LLMs).

This paper delves into the evolution of AI based digital tools and the emergence of LLMs and discusses several key areas to better understand LLMs, the key ethical concerns related to them, and the future of their use in digital writing and beyond. Firstly, we describe and demonstrate the potential that these tools have in creating original, coherent text that can avoid traditional methods of detection by text-matching software. Secondly, we evaluate whether the use of LLM based tools to support students in writing assignments can be considered as plagiarism, academic misconduct, or a breach of academic integrity. Thirdly, we identify the potential these tools have for supporting the education of students, and whether academic staff can detect any such use of

Academic Editors

Section: Special Issue
Senior Editor: Dr Jo-Anne Kelder
Guest Editor: Dr Rebecca Johnke

Publication

Received: 28 November 2022
Revision: 20 December 2022
Accepted: 31 January 2023
Published: 22 February 2023

Copyright: © by the authors, in its year of first publication. This publication is an open access publication under the Creative Commons Attribution [CC BY-ND 4.0](https://creativecommons.org/licenses/by-nd/4.0/) license.

these tools. We end with a discussion on how academic integrity policies of HEIs must be updated to include the recognition of these tools.

By exploring these areas, we contribute to the literature by increasing the awareness amongst practitioners of an emerging technological tool which may be used by students to evade traditional methods of detecting breaches of academic integrity. We also identify key considerations for HEIs as they begin to consider how their academic integrity policies may need to be adjusted to account for the emergence and use of these tools by students and highlight specific areas which may need to be considered. By determining where boundaries may lie in relation to academic integrity, academic misconduct, and plagiarism, we support HEIs in the broader development of their academic integrity policies. Therefore, clarifying how student use of these digital tools may be considered either an acceptable usage of a potentially paradigm shifting educational tool, or an academic integrity violation.

Literature

The Evolution of AI Powered Digital Writing Tools

Students have used digital writing tools to aid them in their assessments since their rise to prominence in the 1980s (Palmquist, 2003), with the development of style and grammar analysis software integrated into word processor technology. These tools have seen a long history of evolution and change, and students believe that these tools are able to support them in improving their writing output (Nobles & Paganucci, 2015). Although digital writing has long been used in HEIs (Kozma, 1991), new technologies mean that there is a constant need to evaluate which tools are available to students, and provide clear guidelines on how these may be used.

Apart from the spelling, grammar, and style checkers built into word processing software, tools have emerged which are designed to go beyond the basics of pointing out more basic errors in composition and provide additional guidance to students on improving their writing. For example, Digital Writing Assistants (DWAs) such as Grammarly, WordTune and Perusall, use Artificial Intelligence (AI) to improve student writing (Fitria, 2021), and are perceived by students to be helpful tools in improving their work (Cavaleri & Dianati, 2016; O'Neill & Russell, 2019). These tools have also demonstrated a particular ability to support English as a Foreign Language (EFL) writers in improving their skills at expressing ideas in written English (Gayed et al., 2022; Nazari et al., 2021; Zhao, 2022), but it is vital that these EFL learners are supported in knowing how to use these responsibly so that they do not accidentally breach any academic integrity policies.

Another category of AI based software tools which have been used by students to support them with their writing are those of Automated Paraphrasing Tools (APTs). APTs are applications which use machine translation to transform one text into another (Rogerson & McCarthy, 2017). These tools were initially developed to support commercial enterprises in search engine optimisation (Q. Zhang et al., 2014) but have been adapted into tools that can be used to support students in the paraphrasing of text. Although the developers behind these tools may make claims that paraphrasing using these tools may not constitute plagiarism (QuillBot, n.d.-b), and technically discourage students from using these tools without correct citation of where original text was taken from (QuillBot, n.d.-a), these claims are disingenuous, given the wide variety of policies that may be in place by various HEIs related to what constitutes AD and plagiarism (Sun, 2013).

Passing the burden of paraphrasing text to a digital tool and adjusting the output to enhance readability and reduce the likelihood that any original text will be spotted by text-matching software is considered to be a case of academic dishonesty and an example of paraphrasing plagiarism (Roe & Perkins, 2022).

Legitimate uses of this software can be again demonstrated in EFL education, where these tools can support students being taught the skill of correct paraphrase (Chen et al., 2015; Park & Yang, 2020; Zhao, 2022). However, it is clear from other reports that students do use these APTs for academically dishonest purposes (Dinneen, 2021), even if this may be unintentional (Prentice & Kinden, 2018). Spotting whether students have used APTs in their work can potentially be identified through the ‘word-salad’ that these tools may produce as an output (Rogerson & McCarthy, 2017), or through technological detection methods which aim to identify machine translated text (Wahle, Ruas, Foltýnek, et al., 2022; Q. Zhang et al., 2014), but this remains an emerging challenge, especially as we consider the next evolution of digital writing tools available to students: LLMs.

The Emergence of Large Language Models

Tools using Artificial Intelligence (AI) to support in the creation of original text have emerged in recent years, in the form of Large Language Models (LLMs) which are able to develop significant amounts of brand-new text based on short input requests. This contrasts with the APT and DWA tools described above, which focus primarily on the manipulation of existing text—paraphrasing, suggesting alterations to text, or providing predictions for sentence completion. These models, trained on broad data sets are also referred to as foundation models (Bommasani et al., 2022).

These tools started to achieve prominence from 2017 with the emergence of a range of transformer-based machine learning models which allow for improved performance on language based tasks when compared to previous machine learning techniques (Vaswani et al., 2017). Of note are the BERT transformer developed by Google in 2018, and the subsequent releases of the Generative Pre-trained Transformer 2 (GPT-2) by OpenAI in November 2019, and the Generative Pre-trained Transformer 3 (GPT-3) in June 2020 (Dale, 2021). These LLMs have received significant attention in popular media (Dale, 2021), including multiple reports of the eventual publication of Alarie and Cockfield’s (2021) journal paper which was generated entirely by GPT-3. There has also been increasing awareness of these tools and how they relate to the future of academic and scientific work as demonstrated by the publication in *Nature* of an overview of the advances in this area (Hutson, 2022). The broad release to the public in November 2022 of ChatGPT by OpenAI marked a significant increase in the ability of the underlying software to create new text using further refined models (GPT-3.5), coupled with a markedly improved user interface (OpenAI, 2022). This release has led to an increased public dialogue on how LLMs may impact academic integrity, with multiple opinion pieces written on this topic (see for example Marche (2022) and Hern (2022)), as well as achieving more than 1 million users within five days of launch (Brockman, 2022).

Integration of LLMs into Commercial Services

Although LLMs can be accessed directly, there is a relatively steep learning curve in understanding how to use the software, and some of the more advanced LLMs such as GPT-3 are currently limited to certain geographic areas and require extensive sign-up procedures.

Because of this, an alternative method that students may use to access the LLMs is through commercial entities who use an Application Programming Interface (API) to connect to the LLMs. These services are targeted at different groups of customers including copywriters, marketers and students and allow user-friendly access to the LLMs. Many of these tools integrate features of DWAs and APTs, and market the product package together as one 'writing support tool' or similar, available as both free and paid for versions.

This inclusion of LLMs into DWAs adds to the potential confusion for students as to whether the use of these tools would be considered legitimate. A subset of these tools is marketed specifically to students to support in the development of full length-assessments, which also include options to generate citations alongside the work. These entities are referred to as Automatic Article Generators (AAGs) by Abd-Elaal et al. (2022) and are an additional method for students to use the underlying LLMs on which they are based. Multiple lists exist online which compare these services against each other, with several of these appearing to be sponsored by some of the services themselves.

Examples

As an example of the fluency of the output that can be created using these tools, the two paragraphs of the section above titled: 'The emergence of Large Language Models' were used as an input to ChatGPT. The prompt to the model was: *'Continue the following text by writing approximately 300 words explaining what ChatGPT is capable of, including examples of how it can be used by different users, with a focus on education. Write in a formal academic style, using in text citations in APA7 style. Provide all references'* The unedited output from this prompt was:

'ChatGPT-3 is a chatbot that is powered by the GPT-3 language model developed by OpenAI. It is a large language model that is trained on a broad dataset and is capable of generating human-like text based on short input requests. The GPT-3 model is trained using transformer-based machine learning techniques, which have been shown to outperform previous machine learning techniques on language-based tasks (Vaswani et al., 2017).

One of the main capabilities of ChatGPT-3 is its ability to generate new text based on short input requests. This means that it can be used to create original content, such as articles, blog posts, and even entire papers. This has significant potential for education, as it could be used to assist students in the writing process, or even to generate complete papers for assignments.

There are a number of potential ways in which ChatGPT-3 could be used in education. For example, it could be used as a writing tool to help students generate ideas for assignments, or to suggest edits and improvements to existing text. It could also be used as a tool for language learning, as it can be trained to generate text in multiple languages.

However, the use of ChatGPT-3 in education also raises some ethical concerns. There are concerns that the use of chatbots like ChatGPT-3 could lead to a decrease in academic integrity, as students may be tempted to use the tool to generate entire papers rather than doing the work themselves. There have been several opinion pieces written on this topic (Marche, 2022; Hern,

2022), and it is important for educators and students to consider the potential consequences of using this technology.

Overall, ChatGPT-3 is a powerful tool that has the potential to revolutionize the way we approach education. It has the ability to generate original text and assist students in the writing process, but it is important to consider the potential ethical implications of its use.

References

- Bommasani, S., Vaswani, A., & Dale, S. (2022). Large Language Models: A Review. arXiv preprint arXiv:2201.11701.
- Dale, S. (2021). The Rise of Large Language Models. *Communications of the ACM*, 64(6), 44-53.
- Hutson, M. (2022). The next step for language models. *Nature*, 588(7840), 395-397.
- Marche, S. (2022). The dangers of AI-generated papers. *The Atlantic*. Retrieved from <https://www.theatlantic.com/ideas/archive/2022/02/ai-generated-papers-threaten-academic-integrity/617370/>
- Hern, A. (2022). Chatbots could undermine academic integrity, warn researchers. *The Guardian*. Retrieved from <https://www.theguardian.com/education/2022/mar/01/chatbots-could-undermine-academic-integrity-warn-researchers>
- OpenAI (2022). ChatGPT: A chatbot powered by GPT-3. Retrieved from <https://openai.com/blog/chatgpt/>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., & Polosukhin, I. (2017). Attention is all you need. In *Advances in neural information processing systems* (pp. 5998-6008).
- (OpenAI, n.d.)

This output was created in approximately 5 seconds at zero cost (ChatGPT is currently in a free research preview stage and costs are likely to change in the near future). The output retained the language style used in the input and expanded on the discussions by providing examples specifically related to education as requested. Although in-text citations and references were provided (including references from the input text), these were all fabricated, with titles, URLs, and journal names being incorrect. However, given that the authors and dates provided in the in-text citations are correct, this may go unnoticed unless thoroughly checked. This does open a potential avenue of detection by academic staff if there is a suspicion that these tools may have been used by students.

A major benefit of ChatGPT as opposed to previous versions of the GPT-3 software is that of the iteration possible through natural language 'discussions' held with the software. For example, in the same conversational 'thread' as the previous output, ChatGPT was asked to: *Continue your response, focusing on how Higher Education Institutions may need to adjust their academic integrity policies based on student use of these technologies. Again, use full APA 7 in-text citations and provide references.* The unedited output created was:

'As the use of ChatGPT-3 and other large language models becomes more widespread, it is likely that Higher Education Institutions (HEIs) will need to adjust their academic integrity policies to take into account the potential use of these technologies by students.

One of the main concerns with the use of chatbots like ChatGPT-3 is the potential for students to use them to generate entire papers or assignments, rather than doing the work themselves (Marche, 2022). This could lead to a decrease in academic integrity, as students may be able to pass off the work of the chatbot as their own. To address this concern, HEIs may need to update their policies to specifically prohibit the use of chatbots or other AI-powered tools for the generation of academic work.

In addition to updating policies, HEIs may also need to invest in tools and technologies to detect the use of chatbots in academic work. Plagiarism detection software, such as Turnitin, can be used to identify text that has been generated by chatbots or other AI tools, but these tools may need to be updated to be able to accurately detect the use of chatbots (Hern, 2022).

Another potential issue with the use of chatbots in education is the potential for the chatbot to produce content that is biased or offensive. Chatbots are trained on large datasets, and if these datasets contain biased or offensive content, the chatbot may produce similar content (Bommasani et al., 2022). To address this concern, HEIs may need to consider implementing policies or guidelines related to the use of chatbots or other AI tools, and providing training to students on the appropriate use of these technologies.

Overall, the use of chatbots like ChatGPT-3 in education raises a number of ethical concerns that HEIs will need to consider when developing or updating their academic integrity policies. It is important for HEIs to carefully consider the potential consequences of the use of these technologies, and to implement policies and guidelines that ensure the integrity of academic work.

References

Bommasani, S., Vaswani, A., & Dale, S. (2022). Large Language Models: A Review. arXiv preprint arXiv:2201.11701.

Marche, S. (2022). The dangers of AI-generated papers. The Atlantic. Retrieved from <https://www.theatlantic.com/ideas/archive/2022/02/ai-generated-papers-threaten-academic-integrity/617370/>

Hern, A. (2022). Chatbots could undermine academic integrity, warn researchers. The Guardian. Retrieved from <https://www.theguardian.com/education/2022/mar/01/chatbots-could-undermine-academic-integrity-warn-researchers>

(OpenAI, n.d.)

Again, the references were falsified and there were some inaccuracies related to the claimed ability of Turnitin to detect AI generated content, but this fluent, five-paragraph output is typical of a response by ChatGPT. It is evident that with some manipulation of the text through iteration using the tool, along with some minor editing or adjustments of references, a student intending to create large blocks of text to include in an assessment could rapidly complete certain tasks with limited effort and costs.

Limitations and Newer LLMs

As demonstrated above, the output produced by the LLMs is fluent, however there are some limitations to the currently available LLMs. Given that the development of these LLMs require pre-training before they are released, they are unable to answer questions or complete tasks which relate to very recent events. As they rely on the content they have been trained with to generate their output, there is the potential for LLMs to produce text that lacks semantic coherence (Dale, 2021), as well as lexical diversity (Gehrmann et al., 2019) through the repetition of expressions (Dehouche, 2021; Fröhling & Zubiaga, 2021).

Although one concern levied against LLMs is that they will generally always provide an answer or output, even if that is factually incorrect or misinterpreted (Dale, 2021; Hutson, 2022), recent research has shown that GPT-3 can be trained to state how confident it is in producing an answer that is factually correct (Lin et al., 2022), which can be used to support users of the LLMs in their use of the tools. ChatGPT in particular can also be challenged if the user believes there are incorrect elements in the output, and can adjust its response accordingly based on any corrections provided by the user.

Recent concerns related to the lack of transparency regarding the models used in currently available LLMs, as well as their computational cost and associated environmental impacts (Sanh et al., 2022; S. Zhang et al., 2022) have resulted in the development of alternative LLMs, including an open source LLM called GPT-J by Narrativa, and a new suite of Open Pre-trained Transformers (OPTs) by Meta. The resulting OPT-175B model by Meta has been shown to have comparable results to GPT-3 in terms of text creation ability, whilst being more open to scrutiny by researchers (S. Zhang et al., 2022) as to the ethical implications of these models; a particularly relevant concern as it relates to the study of academic integrity.

Academic Integrity Considerations of LLMs

Concerns for Academic Integrity

As demonstrated by the examples provided above, the current generation of LLMs are already fluent in their output, and emerging research has suggested that existing LLMs can produce output which humans struggle to identify as being machine created (Abd-Elaal et al., 2022; Clark et al., 2021; Gunser et al., 2021; Köbis & Mossink, 2021; Wahle et al., 2021; Wahle, Ruas, Kirstein, et al., 2022). At the same time, given that the text that is produced by LLMs is uniquely created based on the inputs provided, current research suggests that use of the created text by students is unlikely to be spotted by existing text-matching software tools used by HEIs (Wahle, Ruas, Kirstein, et al., 2022). LLMs therefore represent a clear potential threat to academic integrity as academic staff may be unable to correctly identify the amount of content produced by a student, and therefore provide an accurate evaluation of a student's comprehension and interpretation of the topic at hand.

Definitions

There is a growing body of academic research which is exploring the concerns related to the use of LLMs in student work (Dehouche, 2021; Eaton et al., 2021; Kumar, Mindzak, Eaton, et al., 2022; Kumar, Mindzak, & Racz, 2022; Wilder et al., 2021). A question which has been raised by several of these authors but has yet to be fully answered is whether the use of LLMs by students

can be considered a breach of academic integrity or not. Reaching this determination is challenging as it is dependent on how various HEIs define terms related to academic integrity.

For this paper, we use the definition of academic integrity provided by the Tertiary Education Quality and Standards Agency (TEQSA) of *'the expectation that teachers, students, researchers and all members of the academic community act with: honesty, trust, fairness, respect and responsibility'* (Tertiary Education Quality and Standards Agency [TEQSA], 2020).

Recognising the broadness of such a definition, we regard academic misconduct as a breach of academic integrity principles as specified in policy by a HEI, whether this intentional or not. This clarification is needed given the differences between what HEIs and students might classify as academically dishonest behaviour (Parkinson et al., 2022) and as students can unintentionally breach these policies (Amigud, 2020). A key area of interest related to the use of LLMs is the act of plagiarism which we define as *'misrepresenting the effort that has been carried out by the author of a written document'* (Perkins et al., 2019, p. 5). Although there are existing and well known technological tools to detect simple copy and paste plagiarism by students, we are now operating in an era of more complex 'second generation' (Malesky et al., 2016) plagiarism, which requires ways of thinking to consider what plagiarism really entails, as well as more complex method of detection.

Student Uses of LLMs

Based on these definitions we need to consider a range of possible ways that students may use an LLM to support them in the writing of an assessment. Doing this helps us to understand where we may draw a line between acceptable practice, breaches of academic integrity, academic misconduct and/or plagiarism.

Consider a situation where a student uses an LLM to support them in their work, includes directly copied output of the tools in their submission, and does not state that they used an AI tool. This is similar to a case reported by Dinneen (2021) in reference to student use of an APT, where the student in question was convinced that because they had altered the text of the original author, even through the use of a software tool, this was not plagiarism. However, based on the definitions provided above, given the misrepresentation that have been made, usage of this type could be seen as a relatively clear case of plagiarism by the student (Roe & Perkins, 2022). Student usage of an LLM or LLM based tool to create an entire essay, report, or other assignment without stating where this has come from, or one that is presented in a way that was trying to mislead the reader through false citations, would be even more clearly a case of plagiarism, and likely also a broader breach of academic integrity principles.

However, the picture becomes less clear if the use of LLMs in student writing has been clearly stated (Kumar, Mindzak, Eaton, et al., 2022), if any output text has been modified, edited or enhanced, or if output text has been used as a starting point for the formation of an argument. If additional scholarly work has been carried out by the student, such as the integration of the output text with the use of examples and cited sources, then this complicates the situation further. If the work carried out by a student hasn't been misrepresented, then it does not fall under the strict definition of plagiarism, but it may be considered as academic misconduct depending on the specific policies and regulations in place at individual HEIs. HEIs must therefore be clear in both their policies and practices in clarifying how these tools may be used by students. The difficulty in establishing these policies is even more challenging if we consider the use of LLMs as a tool for student cognitive offloading.

Cognitive Offloading

If student use of LLMs could be considered to be academic misconduct if their use isn't clearly identified, to what extent should we expect a student to state whether they have used LLMs in the creation of their work? It would be unrealistic to expect a student to identify every sentence which has been modified using a spelling or grammar checker, every reference which has been generated by a reference manager, or even any sentences which were completed by the built-in features of DWAs or Microsoft Word's write-ahead function. As Microsoft have signed an exclusive agreement with OpenAI to use GPT-3 and integrate these into Microsoft's Azure platform (Langston, 2021), it is highly likely that LLMs will be integrated more deeply into suites of products that are often provided to students by HEIs in the future (Dale & Viethen, 2021).

Therefore, we may consider LLMs as another potential tool for reducing the cognitive demands required by a task, described by Risko and Gilbert (2016) as a process called cognitive offloading. Dawson (2020) has described how assessment tasks can be developed in a way which makes the use of any such cognitive offloading tools transparent, and a specific usage of LLMs as an example of this is discussed by Dawson (2022, as cited in Sparrow, 2022).

Where then, should this distinction be drawn? Given that many of the more advanced DWAs have multiple features available to students which already claim to integrate AI technologies, it is feasible to believe that a student could very easily and inadvertently breach any policies related to their 'correct' usage, especially when LLMs become even more deeply integrated into DWAs. This already appears to be the case in the example of one DWA called Write Full: a full-service DWA which has integrated paraphrasing tools which may encourage students to use the work of others as inputs, rather than their own writing. Adding LLM based features to the next generation of DWAs to suggest additional points for students to consider, and even offering to write 'suggested' text is a very feasible next step for developers wishing to retain an edge over their competitors in this emerging market.

Plagiarism or Academic Misconduct?

Overall, it is highly challenging to unpick the complexities surrounding student usage of LLMs and tools integrating these, and to make ethical judgements as to whether their use may be considered acceptable or not. Kumar (2022) examines how the use of LLMs may be considered from a range of ethical perspectives, and the broader ethical concerns related to LLMs are considered by Luitse and Denkena (2021), but critical questions remain as to how their use may be encouraged or discouraged on a policy level by HEIs.

Based on the analysis above, we conclude that the use of LLMs or LLM based tools should not be considered as plagiarism or a breach of academic integrity if how these tools have been used is stated clearly and made transparent by students. However, given the significantly more advanced capability of LLM based tools to create new content (as opposed to the refinement of existing content) HEIs must be aware of the potential danger of learning outcomes not being met if LLMs are being used to develop content, rather than this being created by students themselves.

LLM Use Cases in the Classroom: Digital Writing and Beyond

As discussed above, there are significant academic integrity concerns related to how HEIs must adapt to the potential use of these AI tools in student submissions. However, research has highlighted several legitimate use cases of AI technologies integrated into digital tools which can support in the education of students. These include specific use cases by students in writing and

composition classes and in the development of creative outputs, and also how academic staff may integrate these tools in the areas of supporting EFL learners and improving Automated Writing Evaluation (AWE).

Writing and Composition

McKnight (2021) highlights the specific challenges of integrating AI into the teaching of writing and composition, and proposes that composition teachers working with AI should guide students in:

‘.when it is appropriate to incorporate AI in writing projects, what the affordances and constraints of different versions of AI might be, whether and how to acknowledge the role of AI, and what dimensions humans can contribute beyond the efficiency of machines.’
(McKnight, 2021, p. 11).

Using LLMs in the teaching of digital writing therefore becomes a method to support writers in engaging more deeply with a topic, rather than at a surface level. McKnight (2021) discusses how the integration of LLMs into student writing has the potential to encourage new modes of thinking, but that this must be done with consideration as to how AI writing might be used in the future workplace.

Writing inputs and adjusting model parameters for LLMs in a way that results in legible, usable text is a specific technical skill which could also be integrated into writing instruction sessions. Doing so would allow students to improve their meta-linguistic knowledge around a topic (Godwin-Jones, 2022) and therefore improve their potential skills in writing in specific areas.

McKnight (2021) and Anson (2022) both recognise the importance of training students in how LLMs be used as a form of writing co-creation, as well as discussing when this may or may not be appropriate. Taking this approach would require the development of clear policies on both a programme and institutional level to clarify to students exactly how this may be done in student work, and the level of reporting required to avoid any breaches of academic misconduct policies. Limits regarding how much of the end output must be the student’s own work as opposed to that of an AI tool could also be provided to help support students in how to use these tools in an acceptable way.

However, there have already been concerns raised from writers who believe that co-writing with the current generation of LLMs presents significant barriers regarding the control of the writing process (Biermann et al., 2022), and this highlights the further explorations required in the field of human-AI co-creation to better understand the future needs of the writing classroom.

Creative Outputs

LLMs have a demonstrated capability in not only traditional writing activities, but also for creative works such as poetry (Gunser et al., 2021; Köbis & Mossink, 2021) and computer programming (Biderman & Raff, 2022). Other tools also produced by OpenAI include the Dall-E 2 AI system which is able to create and edit detailed digital artwork based on natural language prompts provided by users (OpenAI, n.d.)

As an increasing amount of textual and creative outputs are being co-created by humans and AI based technologies, there is ongoing research exploring how the intricacies of co-creation may be further be codified and explored in the fields of creative writing (M. Lee et al., 2022; Y. Lee et al., 2022), software development (Biderman & Raff, 2022) and artistic outputs (Oppenlaender, 2022). Any such co-creation also raises concerns related to copyright (Dehouche, 2021) and how

these systems can be more transparent, fair, and accountable (Fröhling & Zubiaga, 2021; Oppenlaender, 2022).

Use Cases by Academic Staff to Support Students

Deciding how to identify what is acceptable use of these tools in relation to student submissions of any writing or creative outputs will be a major challenge for HEIs. Using the definitions related to plagiarism and academic integrity discussed above, if the use of AI tools is made clear in any submission of work by a student, the use cases above could not be considered as plagiarism. However, this does highlight again how HEIs must ensure that any academic integrity policies are explicit in how these tools may be used, as opposed to a simple blanket ban on their use, especially considering the broader benefits that these tools may provide in the areas of supporting EFL learners and in AWE.

Supporting EFL Learners

HEIs must also consider how AI based tools are used in language training for EFL students before they enter their main academic programmes. Chen et al., (2015) describe the use of a corpus-based software program similar to an APT to support Chinese students in their study of English. Although this tool does not claim to use AI technology, it introduces students to a digital method which may be used through trial and error to create better textual outputs in their paraphrasing. Gayed et al., (2022) demonstrate how an LLM tool based on GPT-2 can be helpful for EFL students writing English by suggesting predictions which expands on the more basic support offered by existing writing tools. Their study suggested that this tool could support in improving the lexical diversity of student work, but statistically significant results were not obtained.

These studies highlight the specific need for education of students on any policies which discourage the use of these tools on core academic programmes. If students are trained on these tools in pre-university language programmes, and then move into their main studies, it is easy to see that students would expect that they would be able to continue using them. This is especially important given the increased risks of academic misconduct occurring in students with lower English language abilities (Perkins et al., 2018). Therefore, it is clear that academic integrity policies need to be context sensitive (Price, 2002) as students transition throughout their studies and recognise the needs of different groups of students, especially 'international' students (Fatemi & Saito, 2020).

Automated Writing Evaluation (AWE)

Using software tools to support in AWE of written text and providing feedback to students has been demonstrated to be of benefit to writing instructors and students alike (Fu et al., 2022; Strobl et al., 2019; Warschauer & Grimes, 2008). The further development of LLMs may improve this, especially if this integrated into custom tools to support in AWE (Lim et al., 2022), or DWAs such as Grammarly (Godwin-Jones, 2022). Although Grammarly uses its own AI supported tool to correct mistakes in grammar and rather than relying on existing LLM transformers (Grammarly, 2022) the use of LLMs in these tools can particularly reduce the burden on EFL instructors (Godwin-Jones, 2022).

Some of the recognised shortcomings of present AWE tools are the generic nature of the feedback provided (Fu et al., 2022) and an inability to focus feedback on how well a task has been addressed (Allen et al., 2016). The new capabilities of ChatGPT to provide effective critique on creative writing, equivalent to that of human reviewers (Furze, 2022), suggests that these shortcomings may be addressed by LLM based tools.

Can the Use of LLMs be Detected by Academic Staff?

Any actions that HEIs could choose to take against students who use these tools in an unacceptable manner can only be carried out if the outputs of these tools can be identified as such by academic staff responsible for the evaluation of student work. We review the current evidence evaluating how possible this is.

GPT-2 Studies

Abd-Elaal et al., (2022) present a study highlighting the difficulty that academic staff may have in identifying output produced by LLMs, and the role that training may play in supporting in the correct identification of the usage of any such tools. Their results indicate that on average, participants were able to correctly identify sample text as generated by either a human or an LLM at a rate of 59.5%, barely higher than chance alone. Although it is encouraging to note that training of academic staff resulted in an increased ability of academic staff to correctly identify the text sections, the samples were generated using GPT-2 as opposed to GPT-3.

A similar lack of identification ability has been found in creative outputs. Köbis & Mossink (2021) found that GPT-2 created poems could not be reliably identified as such when the authors selected the best outputs created by an LLM, and Gunser et al. (2021) found that poem continuations could not always be correctly identified by professionals with a literature-specific background as machine or human written.

GPT-3 Studies

Studies assessing the ability of GPT-3 produced output show that as more complex LLMs are used, the ability of humans to detect material drops even further. Kumar et al. (2022) presents preliminary work aimed at identifying whether study participants drawn from a range of backgrounds were able to identify whether text was developed by humans or by GPT-3. They found that regardless of background, participants found this to be a challenging task, with a high likelihood of ascribing the AI writing samples to humans. Although this study was not focused on how academic staff may identify machine created output, it highlights the challenges present in both the academy and beyond when it comes to identifying authorship of text.

Clark et al. (2021) also assessed the ability of non-expert evaluators to identify whether text was produced by a human, or a LLM (GPT-2 and GPT-3) but used a large sample (n=780) accessed through the Amazon Mechanical Turk platform. Their results showed that evaluators were able to identify GPT-2 produced text at an accuracy rate of 57.9%, but GPT-3 at a rate of only 49.9%. Providing training to participants using examples of LLM generated text (specifically to demonstrate the ability of LLMs to produce 'creative' output) marginally increased the ability of participants to correctly identify text as machine created. These studies highlight the rapid improvements of LLMs to create text that cannot be easily identified by study participants as machine created.

Methodological Concerns

An unavoidable methodological issue with these experimental studies is that in order to determine whether participants can accurately identify text as machine or human generated, participants need to be aware that some of the text they are about to encounter may be machine generated before participating in an experiment. Given the novelty of these tools, it is likely that even experienced academic staff are simply not aware of the capabilities that these tools have, as demonstrated with the participants in Clark et al.'s study (2021). This may result in their ability to

identify any LLM produced output 'in situ' when evaluating work being even lower than demonstrated in an experimental design.

However, given the limited amount of empirical evidence available, and the potential significant threats that LLMs pose to academic integrity, further research is needed in this area to support HEIs in understanding how academic staff may be trained or supported to detect the use of AI tools in student work.

Technological Methods of Detection

Technological methods to identify text generated by LLMs have been proposed by Gehrmann et al. (2019) who present the results of a study using a tool named GLTR. This study demonstrates that using GLTR can improve the ability of participants to correctly identify detection of machine generated content from 54% to 74%. However, this tool was tested against GPT-2 produced output and used students rather than academic staff as their participants. Solaiman et al. (2019) and Ippolito et al. (2020) also both present tools which showed an encouraging ability in detecting machine created text, although these tools were not tested against the latest generation of LLMs.

Fröhling & Zubiaga (2021) present a promising low-cost detection model which is able to accurately detect machine created text created using GPT-2 and GPT-3, but highlight the ethical challenges of deploying any such detectors which may potentially discriminate against EFL students by incorrectly identifying human created text as machine written—a particular concern in HEIs with a high concentration of non-native English speaking students. While these results do show promise for the potential future possibilities of software to support in the identification of LLM produced output, Bidermann and Raff (2022) have already demonstrated how more advanced models (GPT-J) can fool machine detection programmes such as MOSS. This suggests that tools using more advanced LLMs may be even less detectable by technological means.

Following the increased interest from the general public and academics alike in AI assisted writing after the release of ChatGPT in November 2022, individuals and organisations have either released, or have announced the imminent release of tools which claim to have the ability to detect AI generated text. These tools include GPTZero (<https://gptzero.me/>) and Crossplag AI detect (<https://crossplag.com/ai-content-detector/>) and do show promise in being able to detect the use of AI generated text. However, further study is required to identify the accuracy of these tools, as well as their suitability for use in academic settings to avoid inadvertently accusing students of potential breaches of academic misconduct. Any tools used to support in the machine detection of LLM output must be continually re-evaluated as new LLMs emerge, as well as methods to avoid detection of any tools are developed, resulting in an ongoing 'arms race' scenario (Roe & Perkins, 2022).

Given that both academic staff, as well as technological methods of detection are unable to accurately detect machine generated text and therefore student uses of LLM based tools, this presents a clear threat to academic integrity for HEIs, requiring a range of adjustments to be made in both practice and policy.

Academic Integrity Policy Adjustments

Although a certain amount of cognitive offloading using DWAs and other digital tools may be considered acceptable by HEIs, there needs to be careful development of the academic integrity policies of HEIs to clarify how LLMs based tools may be used by students, and whether their use may be considered as academic misconduct.

Academic integrity policies must recognise these tools by name to make it clear to students and staff that the HEI is aware of this technology and how it may be used by students and staff. Any limits to their use, or statements which students should provide if these tools are used should be stated and unambiguous. Specific examples of acceptable and unacceptable usage should be provided to improve the understanding of students, as well as staff involved in any academic misconduct decisions. The final policy should also be widely communicated to students and staff, with training given to support whichever approach is taken.

Given that existing research has demonstrated that academic staff are unlikely to be able to accurately identify the usage of LLMs in student work, a blanket ban of these tools is likely unenforceable. This lack of enforceability, coupled with the clear benefits that these tools may bring to the education of certain groups of students, means that an approach which seeks to completely ban the usage of LLM based tools is therefore not recommended. A policy approach which may be more suitable could instead have a more nuanced approach which recognises the potential benefits that LLMs may bring, the evolving social understanding of plagiarism, and the changing nature of digital writing and human-AI co-creation.

Conclusion

What we as academic staff define as plagiarism is rapidly changing due to the social construction of this term (Anson, 2022). Although evidence suggests that plagiarism overall has decreased between 1990 and 2020 (Curtis, 2022), the COVID-19 pandemic has caused an increase in both the detected number of academic dishonesty cases (Henderson et al., 2022; Jenkins et al., 2022; Lancaster & Cotarlan, 2021), as well as increases in student or academic staff perceptions of academic dishonesty occurring (Amzalag et al., 2021; Reedy et al., 2021; Walsh et al., 2021). This paper has explored a rapidly developing new category of digital tools which are of concern if we wish to maintain academic integrity in a post-pandemic world: AI based Large Language Models and associated tools and software.

We have identified how LLMs can produce coherent, original text that students may potentially use in assessments, and conclude that LLMs have already progressed to the point that neither trained academic staff or technological tools can consistently determine whether text is generated by an LLM or by a human.

Deciding whether any particular use of LLMs by students may be defined as academic misconduct will be determined by the future policies of any given HEI, and this highlights the importance of creating clear academic integrity policies and educating students in any acceptable use cases of LLMs. Although we have identified the potential ways that LLMs can support in digital writing and beyond, the inability to accurately detect whether LLMs have been used presents a clear threat to the academic integrity of HEIs. However, it is not the use of the tools themselves that defines whether plagiarism or a breach of academic integrity has occurred, but whether any such use is made clear.

Although it is unlikely that AI will ever take over the traditional role of a teacher (Cope et al., 2021), the future development of LLMs and broader AI supported digital tools have a strong potential for improving the experiences of students and teachers alike in the next generation of HEI classrooms, both in writing instruction and beyond. This view is echoed by Godwin-Jones (2022) who highlight the possible co-creation role that these systems may have to play in future educational scenarios. It is clear from the studies presented in this paper that there are significant challenges still left to address in this rapidly developing area. Given that the use of the current

generation of LLMs cannot be accurately detected by academic staff or technical means of detection, the likelihood of accurately detecting any usage of these tools by students in their submissions typical academic will likely not improve and may even decrease further as new LLMs are developed.

This situation, coupled with the identified difficulties in determining whether we can even consider the use of such tools to be a breach of academic integrity, and the potential benefits of LLM based tools means that a blanket ban of these tools at an institutional level is neither feasible, nor enforceable. Despite the potential threats to academic integrity presented in the paper, we believe that the future integration of LLMs and other AI supported digital tools into the classroom environment is highly likely, and therefore HEIs must consider the implications of this in future policy development.

References

- Abd-Elaal, E.-S., Gamage, S. H., & Mills, J. E. (2022). Assisting academics to identify computer generated writing. *European Journal of Engineering Education*, 1–21.
<https://doi.org/10.1080/03043797.2022.2046709>
- Alarie, B., & Cockfield, A. (2021). Will machines replace us?: Machine-authored texts and the future of scholarship. *Law, Technology and Humans*, 3(2), 5–11.
<https://doi.org/10.5204/lthj.2089>
- Allen, L. K., Jacovina, M. E., & McNamara, D. S. (2016). Computer-Based Writing Instruction. In *Handbook of writing research* (Second edition, pp. 316–329). Guildford.
- Amigud, A. (2020). Cheaters on Twitter: An analysis of engagement approaches of contract cheating services. *Studies in Higher Education*, 45(3), 692–705.
<https://doi.org/10.1080/03075079.2018.1564258>
- Amzalag, M., Shapira, N., & Dolev, N. (2021). Two Sides of the Coin: Lack of Academic Integrity in Exams During the Corona Pandemic, Students' and Lecturers' Perceptions. *Journal of Academic Ethics*. <https://doi.org/10.1007/s10805-021-09413-5>
- Anson, C. M. (2022). AI-Based Text Generation and the Social Construction of “Fraudulent Authorship”: A Revisitation. *Composition Studies*, 50, 37–46.
<https://compositionstudiesjournal.files.wordpress.com/2022/07/anson.pdf>
- Biderman, S., & Raff, E. (2022). *Fooling MOSS Detection with Pretrained Language Models* (arXiv:2201.07406). arXiv. <https://doi.org/10.48550/arXiv.2201.07406>
- Biermann, O. C., Ma, N. F., & Yoon, D. (2022). From Tool to Companion: Storywriters Want AI Writers to Respect Their Personal Values and Writing Strategies. *Designing Interactive Systems Conference*, 1209–1227. <https://doi.org/10.1145/3532106.3533506>
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., ... Liang, P. (2022). *On the Opportunities and Risks of Foundation Models* (arXiv:2108.07258). arXiv.
<http://arxiv.org/abs/2108.07258>
- Brockman, G. [@gdb]. (2022, December 5). *ChatGPT just crossed 1 million users; it's been 5 days since launch*. [Tweet]. Twitter.
<https://twitter.com/gdb/status/1599683104142430208>
- Cavaleri, M. R., & Dianati, S. (2016). You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students. *Journal of Academic Language and Learning*, 10(1), Article 1.
<https://journal.aall.org.au/index.php/jall/article/view/393>

- Chen, M.-H., Huang, S.-T., Chang, J. S., & Liou, H.-C. (2015). Developing a corpus-based paraphrase tool to improve EFL learners' writing skills. *Computer Assisted Language Learning*, 28(1), 22–40. <https://doi.org/10.1080/09588221.2013.783873>
- Clark, E., August, T., Serrano, S., Haduong, N., Gururangan, S., & Smith, N. A. (2021). All That's `Human' Is Not Gold: Evaluating Human Evaluation of Generated Text. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 7282–7296. <https://doi.org/10.18653/v1/2021.acl-long.565>
- Clarke, O., Chan, W. Y. D., Bukuru, S., Logan, J., & Wong, R. (2022). Assessing knowledge of and attitudes towards plagiarism and ability to recognize plagiaristic writing among university students in Rwanda. *Higher Education*. <https://doi.org/10.1007/s10734-022-00830-y>
- Cope, B., Kalantzis, M., & Searsmith, D. (2021). Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies. *Educational Philosophy and Theory*, 53(12), 1229–1245. <https://doi.org/10.1080/00131857.2020.1728732>
- Curtis, G. J. (2022). Trends in plagiarism and cheating prevalence: 1990–2020 and beyond. In D. Rettinger & T. Bertram Gallant (Eds.), *Cheating academic integrity* (pp. 11–44). Jossey-Bass. <https://www.wiley.com/en-au/Cheating+Academic+Integrity%3A+Lessons+from+30+Years+of+Research-p-9781119868170>
- Dale, R. (2021). GPT-3: What's it good for? *Natural Language Engineering*, 27(1), 113–118. <https://doi.org/10.1017/S1351324920000601>
- Dale, R., & Viethen, J. (2021). The automated writing assistance landscape in 2021. *Natural Language Engineering*, 27(4), 511–518. <https://doi.org/10.1017/S1351324921000164>
- Dawson, P. (2020). Cognitive Offloading and Assessment. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining University Assessment in a Digital World* (pp. 37–48). Springer International Publishing. https://doi.org/10.1007/978-3-030-41956-1_4
- Dehouche, N. (2021). Plagiarism in the age of massive Generative Pre-trained Transformers (GPT-3). *Ethics in Science and Environmental Politics*, 21, 17–23. <https://doi.org/10.3354/ese00195>
- Dinneen, C. (2021). Students' use of digital translation and paraphrasing tools in written assignments on Direct Entry English Programs. *English Australia Journal*, 37(1), 40–51. <https://doi.org/10.3316/informit.748262877947586>

- Eaton, S. E., Mindzak, M., & Morrison, R. (2021, June 29). *Artificial Intelligence, Algorithmic Writing & Educational Ethics*. Canadian Society for the Study of Education, Edmonton, AB, Canada. Werklund School of Education. <https://tinyurl.com/222yctbb>
- Fatemi, G., & Saito, E. (2020). Unintentional plagiarism and academic integrity: The challenges and needs of postgraduate international students in Australia. *Journal of Further and Higher Education*, 44(10), 1305–1319. <https://doi.org/10.1080/0309877X.2019.1683521>
- Fitria, T. N. (2021). Grammarly as AI-powered English Writing Assistant: Students' Alternative for Writing English. *Metathesis: Journal of English Language, Literature, and Teaching*, 5(1), Article 1. <https://doi.org/10.31002/metathesis.v5i1.3519>
- Fröhling, L., & Zubiaga, A. (2021). Feature-based detection of automated language models: Tackling GPT-2, GPT-3 and Grover. *PeerJ Computer Science*, 7, e443. <https://doi.org/10.7717/peerj-cs.443>
- Fu, Q.-K., Zou, D., Xie, H., & Cheng, G. (2022). A review of AWE feedback: Types, learning outcomes, and implications. *Computer Assisted Language Learning*, 0(0), 1–43. <https://doi.org/10.1080/09588221.2022.2033787>
- Furze, L. (2022, December 6). Can an AI critique human writing? *Leon Furze*. <https://leonfurze.com/2022/12/06/can-an-ai-critique-human-writing/>
- Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing Assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, 3, 100055. <https://doi.org/10.1016/j.caeai.2022.100055>
- Gehrmann, S., Strobelt, H., & Rush, A. M. (2019). *GLTR: Statistical Detection and Visualization of Generated Text* (arXiv:1906.04043). arXiv. <https://doi.org/10.48550/arXiv.1906.04043>
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, 26(2), 5–24. <https://doi.org/10.125/73474>
- Grammarly. (2022, August 29). *Innovating the Basics: Achieving Superior Precision and Recall in Grammatical Error Correction*. Grammarly Engineering Blog. <https://www.grammarly.com/blog/engineering/innovating-the-basics/>
- Grijalva, T. C., Kerkvliet, J., & Nowell, C. (2006). Academic honesty and online courses. *College Student Journal*, 40(1), Article 1. <https://tinyurl.com/3beej9bb>
- Gunser, V. E., Gottschling, S., Brucker, B., Richter, S., & Gerjets, P. (2021). Can Users Distinguish Narrative Texts Written by an Artificial Intelligence Writing Tool from Purely Human Text? *International Conference on Human-Computer Interaction*, 520–527. https://doi.org/10.1007/978-3-030-78635-9_67

- Henderson, M., Chung, J., Awdry, R., Mundy, M., Bryant, M., Ashford, C., & Ryan, K. (2022). Factors associated with online examination cheating. *Assessment & Evaluation in Higher Education*, 0(0), 1–15. <https://doi.org/10.1080/02602938.2022.2144802>
- Hern, A. (2022, December 4). AI bot ChatGPT stuns academics with essay-writing skills and usability. *The Guardian*. <https://www.theguardian.com/technology/2022/dec/04/ai-bot-chatgpt-stuns-academics-with-essay-writing-skills-and-usability>
- Hutson, M. (2022). Could AI help you to write your next paper? *Nature*, 611(7934), 192–193. <https://doi.org/10.1038/d41586-022-03479-w>
- Ippolito, D., Duckworth, D., Callison-Burch, C., & Eck, D. (2020). *Automatic Detection of Generated Text is Easiest when Humans are Fooled* (arXiv:1911.00650). arXiv. <https://doi.org/10.48550/arXiv.1911.00650>
- Jenkins, B. D., Golding, J. M., Le Grand, A. M., Levi, M. M., & Pals, A. M. (2022). When Opportunity Knocks: College Students' Cheating Amid the COVID-19 Pandemic. *Teaching of Psychology*, 00986283211059067. <https://doi.org/10.1177/00986283211059067>
- Kaqinari, T., Makarova, E., Audran, J., Döring, A. K., Göbel, K., & Kern, D. (2021). The switch to online teaching during the first COVID-19 lockdown: A comparative study at four European universities. *Journal of University Teaching & Learning Practice*, 18(5). <https://doi.org/10.53761/1.18.5.10>
- Köbis, N., & Mossink, L. D. (2021). Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry. *Computers in Human Behavior*, 114, 106553. <https://doi.org/10.1016/j.chb.2020.106553>
- Kozma, R. B. (1991). Computer-based writing tools and the cognitive needs of novice writers. *Computers and Composition*, 8(2), 31–45. [https://doi.org/10.1016/8755-4615\(91\)80046-G](https://doi.org/10.1016/8755-4615(91)80046-G)
- Kumar, R., Mindzak, M., Eaton, S. E., & Morrison, R. (2022). *AI & AI: Exploring the contemporary intersections of artificial intelligence and academic integrity*. Canadian Society for the Study of Higher Education Annual Conference, Online. Werklund School of Education. <https://tinyurl.com/ycknz8fd>
- Kumar, R., Mindzak, M., & Racz, R. (2022). *Who Wrote This? The Use of Artificial Intelligence in the Academy*. <http://hdl.handle.net/10464/16532>
- Lancaster, T., & Cotarlan, C. (2021). Contract cheating by STEM students through a file sharing website: A Covid-19 pandemic perspective. *International Journal for Educational Integrity*, 17(1), Article 1. <https://doi.org/10.1007/s40979-021-00070-0>

- Langston, J. (2021, November 2). *New Azure OpenAI Service combines access to powerful GPT-3 language models with Azure's enterprise capabilities*. The AI Blog. <https://blogs.microsoft.com/ai/new-azure-openai-service/>
- Lanier, M. M. (2006). Academic Integrity and Distance Learning. *Journal of Criminal Justice Education*, 17(2), Article 2. <https://doi.org/10.1080/10511250600866166>
- Lee, M., Liang, P., & Yang, Q. (2022). CoAuthor: Designing a Human-AI Collaborative Writing Dataset for Exploring Language Model Capabilities. *CHI Conference on Human Factors in Computing Systems*, 1–19. <https://doi.org/10.1145/3491102.3502030>
- Lee, Y., Kim, T. S., Chang, M., & Kim, J. (2022). Interactive Children's Story Rewriting Through Parent-Children Interaction. *Proceedings of the First Workshop on Intelligent and Interactive Writing Assistants (In2Writing 2022)*, 62–71. <https://doi.org/10.18653/v1/2022.in2writing-1.9>
- Lim, K., Song, J., & Park, J. (2022). Neural automated writing evaluation for Korean L2 writing. *Natural Language Engineering*, 1–23. <https://doi.org/10.1017/S1351324922000298>
- Lin, S., Hilton, J., & Evans, O. (2022). *Teaching Models to Express Their Uncertainty in Words* (arXiv:2205.14334). arXiv. <https://doi.org/10.48550/arXiv.2205.14334>
- Luitse, D., & Denkena, W. (2021). The great Transformer: Examining the role of large language models in the political economy of AI. *Big Data & Society*, 8(2), 20539517211047736. <https://doi.org/10.1177/20539517211047734>
- Malesky, L. A., Baley, J., & Crow, R. (2016). Academic Dishonesty: Assessing the Threat of Cheating Companies to Online Education. *College Teaching*, 64(4), 178–183. <https://doi.org/10.1080/87567555.2015.1133558>
- Marche, S. (2022, December 6). *The College Essay Is Dead*. The Atlantic. <https://www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371/>
- McKnight, L. (2021). Electric Sheep? Humans, Robots, Artificial Intelligence, and the Future of Writing. *Changing English*, 28(4), 442–455. <https://doi.org/10.1080/1358684X.2021.1941768>
- Miller, A., & Young-Jones, A. D. (2012). Academic integrity: Online classes compared to face-to-face classes. *Journal of Instructional Psychology*, 39(3–4), Article 3–4. <https://tinyurl.com/mrxnp9zy>
- Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: Randomized controlled trial. *Heliyon*, 7(5), e07014. <https://doi.org/10.1016/j.heliyon.2021.e07014>

- Nobles, S., & Paganucci, L. (2015). Do Digital Writing Tools Deliver? Student Perceptions of Writing Quality Using Digital Tools and Online Writing Environments. *Computers and Composition*, 38, 16–31. <https://doi.org/10.1016/j.compcom.2015.09.001>
- ONeill, R., & Russell, A. (2019). Stop! Grammar time: University students' perceptions of the automated feedback program Grammarly. *Australasian Journal of Educational Technology*, 35(1), Article 1. <https://doi.org/10.14742/ajet.3795>
- OpenAI. (n.d.). *ChatGPT*. Retrieved 21 December 2022, from <https://chat.openai.com>
- OpenAI. (n.d.). *DALL·E 2*. DALL·E 2. Retrieved 25 November 2022, from <https://openai.com/dall-e-2/>
- OpenAI. (2022, November 30). *ChatGPT: Optimizing Language Models for Dialogue*. OpenAI. <https://openai.com/blog/chatgpt/>
- Oppenlaender, J. (2022). *A Taxonomy of Prompt Modifiers for Text-To-Image Generation* (arXiv:2204.13988). arXiv. <https://doi.org/10.48550/arXiv.2204.13988>
- Palmquist, M. (2003). A brief history of computer support for writing centers and writing-across-the-curriculum programs. *Computers and Composition*, 20(4), 395–413. <https://doi.org/10.1016/j.compcom.2003.08.013>
- Park, J.-H., & Yang, I. Y. (2020). Utilizing an AI-Based Grammar Checker in an EFL Writing Classroom. *Academic Journal of Applied Linguistics*, 36(1), 97–120. <https://doi.org/10.17154/kjal.2020.3.36.1.97>
- Parkinson, A. L., Hatje, E., Kynn, M., Kuballa, A. V., Donkin, R., & Reinke, N. B. (2022). Collusion is still a tricky topic: Student perspectives of academic integrity using assessment-specific examples in a science subject. *Assessment & Evaluation in Higher Education*, 47(8), 1416–1428. <https://doi.org/10.1080/02602938.2022.2040947>
- Perkins, M., Basar Gezgin, U., & Gordon, R. (2019). Plagiarism in higher education: Classification, causes and controls. *Pan-Pacific Management Science*, 2, 3–21. <https://doi.org/10.13140/RG.2.2.20694.11841>
- Perkins, M., Gezgin, U. B., & Roe, J. (2018). Understanding the Relationship between Language Ability and Plagiarism in Non-Native English Speaking Business Students. *Journal of Academic Ethics*, 16(4), Article 4. <https://doi.org/10.1007/s10805-018-9311-8>
- Prentice, F. M., & Kinden, C. E. (2018). Paraphrasing tools, language translation tools and plagiarism: An exploratory study. *International Journal for Educational Integrity*, 14(1), 11. <https://doi.org/10.1007/s40979-018-0036-7>
- Price, M. (2002). Beyond 'Gotcha!': Situating Plagiarism in Policy and Pedagogy. *College Composition and Communication*, 54(1), 88–115. <https://doi.org/10.2307/1512103>

- QuillBot. (n.d.-a). *Can I use QuillBot for plagiarism?* QuillBot Help Center. Retrieved 16 November 2022, from <https://help.quillbot.com/hc/en-us/articles/1500007485841-Can-I-use-QuillBot-for-plagiarism->
- QuillBot. (n.d.-b). *Is using a paraphrasing tool plagiarizing?* QuillBot Help Center. Retrieved 16 November 2022, from <https://help.quillbot.com/hc/en-us/articles/4408075614103-Is-using-a-paraphrasing-tool-plagiarizing->
- Rapanta, C., Botturi, L., Goodyear, P., Guàrdia, L., & Koole, M. (2021). Balancing Technology, Pedagogy and the New Normal: Post-pandemic Challenges for Higher Education. *Postdigital Science and Education*, 3(3), 715–742. <https://doi.org/10.1007/s42438-021-00249-1>
- Reedy, A., Pfitzner, D., Rook, L., & Ellis, L. (2021). Responding to the COVID-19 emergency: Student and academic staff perceptions of academic integrity in the transition to online exams at three Australian universities. *International Journal for Educational Integrity*, 17(1), Article 1. <https://doi.org/10.1007/s40979-021-00075-9>
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive Offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. <https://doi.org/10.1016/j.tics.2016.07.002>
- Roe, J., & Perkins, M. (2022). What are Automated Paraphrasing Tools and how do we address them? A review of a growing threat to academic integrity. *International Journal for Educational Integrity*, 18(1), Article 1. <https://doi.org/10.1007/s40979-022-00109-w>
- Rogerson, A. M., & McCarthy, G. (2017). Using Internet based paraphrasing tools: Original work, patchwriting or facilitated plagiarism? *International Journal for Educational Integrity*, 13(1), 2. <https://doi.org/10.1007/s40979-016-0013-y>
- Sanh, V., Webson, A., Raffel, C., Bach, S. H., Sutawika, L., Alyafeai, Z., Chaffin, A., Stiegler, A., Scao, T. L., Raja, A., Dey, M., Bari, M. S., Xu, C., Thakker, U., Sharma, S. S., Szczechla, E., Kim, T., Chhablani, G., Nayak, N., ... Rush, A. M. (2022). *Multitask Prompted Training Enables Zero-Shot Task Generalization* (arXiv:2110.08207). arXiv. <https://doi.org/10.48550/arXiv.2110.08207>
- Solaiman, I., Brundage, M., Clark, J., Askill, A., Herbert-Voss, A., Wu, J., Radford, A., Krueger, G., Kim, J. W., Kreps, S., McCain, M., Newhouse, A., Blazakis, J., McGuffie, K., & Wang, J. (2019). *Release Strategies and the Social Impacts of Language Models* (arXiv:1908.09203). arXiv. <https://doi.org/10.48550/arXiv.1908.09203>
- Sparrow, J. (2022, November 18). 'Full-on robot writing': The artificial intelligence challenge facing universities. *The Guardian*. <https://www.theguardian.com/australia-news/2022/nov/19/full-on-robot-writing-the-artificial-intelligence-challenge-facing-universities>

- St-Onge, C., Ouellet, K., Lakhal, S., Dubé, T., & Marceau, M. (2022). COVID-19 as the tipping point for integrating e-assessment in higher education practices. *British Journal of Educational Technology*, 53(2), Article 2. <https://doi.org/10.1111/bjet.13169>
- Strobl, C., Ailhaud, E., Benetos, K., Devitt, A., Kruse, O., Proske, A., & Rapp, C. (2019). Digital support for academic writing: A review of technologies and pedagogies. *Computers & Education*, 131, 33–48. <https://doi.org/10.1016/j.compedu.2018.12.005>
- Stuber-McEwen, D., Wiseley, P., & Hoggatt, S. (2009). Point, click, and cheat: Frequency and type of academic dishonesty in the virtual classroom. *Online Journal of Distance Learning Administration*, 12(3), Article 3. <https://tinyurl.com/2awz7b7y>
- Sun, Y.-C. (2013). Do journal authors plagiarize? Using plagiarism detection software to uncover matching text across disciplines. *Journal of English for Academic Purposes*, 12(4), 264–272. <https://doi.org/10.1016/j.jeap.2013.07.002>
- Tertiary Education Quality and Standards Agency. (2020, October 28). *What is academic integrity?* Tertiary Education Quality and Standards Agency. <https://www.teqsa.gov.au/what-academic-integrity>
- Vargo, D., Zhu, L., Benwell, B., & Yan, Z. (2021). Digital technology use during COVID-19 pandemic: A rapid review. *Human Behavior and Emerging Technologies*, 3(1), 13–24. <https://doi.org/10.1002/hbe2.242>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems*, 30. <https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>
- Wahle, J. P., Ruas, T., Foltýnek, T., Meuschke, N., & Gipp, B. (2022). Identifying Machine-Paraphrased Plagiarism. In M. Smits (Ed.), *Information for a Better World: Shaping the Global Future* (pp. 393–413). Springer International Publishing. https://doi.org/10.1007/978-3-030-96957-8_34
- Wahle, J. P., Ruas, T., Kirstein, F., & Gipp, B. (2022). How Large Language Models are Transforming Machine-Paraphrased Plagiarism. *ArXiv Preprint ArXiv:2210.03568*. <https://arxiv.org/pdf/2210.03568.pdf>
- Wahle, J. P., Ruas, T., Meuschke, N., & Gipp, B. (2021). Are neural language models good plagiarists? A benchmark for neural paraphrase detection. *2021 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*, 226–229. <https://doi.org/10.1109/JCDL52503.2021.00065>
- Walsh, L. L., Lichti, D. A., Zambrano-Varghese, C. M., Borgaonkar, A. D., Sodhi, J. S., Moon, S., Wester, E. R., & Callis-Duehl, K. L. (2021). Why and how science students in the

- United States think their peers cheat more frequently online: Perspectives during the COVID-19 pandemic. *International Journal for Educational Integrity*, 17(1), Article 1. <https://doi.org/10.1007/s40979-021-00089-3>
- Warschauer, M., & Grimes, D. (2008). Automated Writing Assessment in the Classroom. *Pedagogies: An International Journal*, 3(1), 22–36. <https://doi.org/10.1080/15544800701771580>
- Watson, G., & Sottile, J. (2010). Cheating in the Digital Age: Do Students Cheat More in Online Courses? *Online Journal of Distance Learning Administration*, 13(1). <https://tinyurl.com/yckenh4f>
- Wilder, N., Weßels, D., Gröpler, J., Klein, A., & Mundorf, M. (2021). Who is Responsible for Integrity in the Age of Artificial Intelligence? An Analysis Using the Example of Academic Writing. *European Conference on Academic Integrity and Plagiarism 2021*, 179–181. <https://tinyurl.com/2s4r5dyr>
- Zhang, Q., Wang, D. Y., & Voelker, G. M. (2014). DSpin: Detecting Automatically Spun Content on the Web. *Proceedings 2014 Network and Distributed System Security Symposium*. Network and Distributed System Security Symposium, San Diego, CA. <https://doi.org/10.14722/ndss.2014.23004>
- Zhang, S., Roller, S., Goyal, N., Artetxe, M., Chen, M., Chen, S., Dewan, C., Diab, M., Li, X., Lin, X. V., Mihaylov, T., Ott, M., Shleifer, S., Shuster, K., Simig, D., Koura, P. S., Sridhar, A., Wang, T., & Zettlemoyer, L. (2022). *OPT: Open Pre-trained Transformer Language Models* (arXiv:2205.01068). arXiv. <https://doi.org/10.48550/arXiv.2205.01068>
- Zhao, X. (2022). Leveraging Artificial Intelligence (AI) Technology for English Writing: Introducing Wordtune as a Digital Writing Assistant for EFL Writers. *RELC Journal*. <https://doi.org/10.1177/00336882221094089>