Journal of Intellectual Property Rights
Vol 31, November 2026, pp. 898-907
DOI: 10.56042/jipr.v31i6.19742

AI Plagiarism Policy for Indiain the Age of ChatGPT

Paarth Naithani[†]

Jindal Global Law School, O.P. Jindal University, Haryana - 131 001, India

Received: 15th May 2025; revised: 25th September 2025

In this age where AI tools such as ChatGPT are used tosubstituteresearch andwriting, how should plagiarism policies of academic institutions ensure honesty and scientific progress? This paper approaches the question by considering the concept of plagiarism and providing insights into students' progression in adjusting with plagiarism software from the years pre-ChatGPT to years post introduction of ChatGPT. Considering the discovered practicalities, the paper then proposes an AI plagiarism taxonomy. The taxonomy of AI plagiarism classifies the following kinds of plagiarism using AI - verbatim copying plagiarism, uncited paraphrasing plagiarism, AI idea plagiarism, language change plagiarism, misrepresented citation AI plagiarism, and prohibited use of AI plagiarism. The paper then also proposes a two-tiered AI plagiarism policy where the first tier follows the UGC Plagiarism Regulations, 2018. The second tier recognises the need for human oversight and possible software development to flag (1) sentences which are stating a fact or idea that is not common knowledge and where citation is expected (2) sentences which are stating a fact or idea which has been cited; however, the citation is misrepresenting the source and (3)missing acknowledgement of use of AI in idea generation, text generation, paraphrasing and other uses. Therefore, the paper proposes a progressive addition to the traditional understanding of plagiarism as textual similarity to a broader, idea- and citation-based understanding of academic integrity.

Keywords: Large Language Models, ChatGPT, Plagiarism, Turnitin, Plagiarism Software, Paraphrasing

Plagiarism policies of today face challenges posed by LLMs such as ChatGPT. There are ways of bypassing the plagiarism software to pass text as non plagiarised when the text has not, in fact, been written by the person themselves.[1] Thus, there is a need for reconsidering policies to protect academic progress and honesty in the synthetic society[2], i.e. the age where it is difficult to distinguish whether AI has been used for writing a text instead of humans having written it themselves.

This paper first discusses the literature which is divided into the following themes: the idea and kinds of plagiarism – stance of global universities and publishers, plagiarism as an academic offence in India, and copyright infringement vis-a-vis plagiarism and the role of AI. The paper then discusses the research question and methodology and progresses to engage in describing the results. The paper develops a policy proposal which is an AI plagiarism taxonomy and a two-tiered AI plagiarism policy where the first tier follows the UGC Plagiarism Regulations, 2018.[3] The second tier is proposed to realistically integrate human oversight in AI plagiarism detection.

[†]Corresponding author: Email: pnaithani@jgu.edu.in

The Idea and Kinds of plagiarism –Stance of Global Universities and Publishers

The idea of plagiarism and the different kinds of plagiarism have been explained by a Harvard University guide.[4] The University guide states that "it is considered plagiarism to draw any idea or any language from someone else without adequately crediting the source."[4] It is important to avoid plagiarism because there is a responsibility for careful and original thought, and to represent others' ideas accurately while oneself contributing to the discussion.[4]

Plagiarism can be verbatim plagiarism where one copies word for word from another source, inadequate paraphrase plagiarism where an idea is not completely restated in ones' own words and the language is too close to the original, uncited paraphrase plagiarism where somebody else's' ideas are paraphrased without citation, uncited quotation plagiarism where a source material is put in quotations but not cited, and plagiarism where material is used from someone else's' work especially when people work in groups.[4] There are subtler forms of plagiarism where source is misrepresented and those papers are ignored which are found at a later stage of ones' research where a same or similar argument has already been made.[4]

However, there are exceptions to plagiarism which are use of common knowledge which is not attributable to one source.[4] But these should not be facts, ideas or language that are distinct to somebody's work.[4]Thus, widely known scientific or historical facts also do not need to be cited.[4]

The Oxford University definition of plagiarism can be referred to understand more facets of the plagiarism. Plagiarism is "[p]resenting work or ideas from another source as your own, with or without consent of the original author, by incorporating it into your work without full acknowledgement. All published and unpublished material, whether in manuscript, printed or electronic form, is covered under this definition, as is the use of material generated wholly or in part through use of artificial intelligence (save when use of AI for assessment has received prior authorisation e.g. as a reasonable adjustment for a student's disability)."[5]

In the context of use of AI in ones' work, it is useful to consider the stance of some global publishers on AI use in manuscripts.

The stance of different publishers indicates that use of AI in publishing comes with several challenges. (The stance of the different publishers is mentioned in Table 1 and the readers can continue now to first read the table and then return to the text here.) WIPO has alsorecognized challenges with generative AI and stated the importance of using plagiarism checkers to check for infringement before using output.[11]While there has been no legislative intervention or judicial decision on plagiarism and AI use in India, a student had unsuccessfully approached Punjab and Haryana High Court challenging the decision of a University Unfair Means Committee which failed him because the Turnitin report showed his submission as 88% AI generated.[12] While India's University Grants Commission (UGC) does not have a specific plagiarism policy for AI, the stance of UGC on plagiarism can be discussed.

Plagiarism as an Academic Offence in India – UGC Regulations

In India, there are the University Grants Commission (Promotion of Academic Integrity and Prevention of Plagiarism in Higher Educational Institutions) Regulations, 2018.[3] The regulation defines plagiarism in Rule 2(l) as "the practice of taking someone else's work or idea and passing them as one's own."[3] There are penalties prescribed by Rule 12.1 when there is plagiarism in the submission of thesis and dissertations. For Level 0 which is similarities upto 10%, there is no penalty. For Level 1 which is similarities above 10%-40%, a revised script must be submitted within a stipulated time not exceeding 6 months. For Level 2 which is similarities above 40%-60%, the student shall be debarred from submitting a revised script until 1 year. For level 3 which is similarities above 60%, the students' program registration would be cancelled. The notes to Rule 12.1 indicate that when a student repeatedly engages in plagiarism, they shall be punished one level higher than the previous level committed by them. The note also clarifies that if plagiarism is proved after the award of degree or credit, such

Table 1 — Stance of publishers on AI use in manuscripts

Publisher	Stance on AI use in manuscripts
Sage	"The use of AI tools that can produce content such as generating references, text, images or any other form of content must be disclosed when used by authors or reviewers. Authors should cite original sources, rather than Generative AI tools as primary sources within the references. If your submission was primarily or partially generated using AI, this must be disclosed upon submission so the Editorial team can evaluate the content generated."[6]
Springer	"Large Language Models (LLMs), such as ChatGPT, do not currently satisfy our authorship criteria. Notably an attribution of authorship carries with it accountability for the work, which cannot be effectively applied to LLMs. Use of an LLM should be properly documented in the Methods section (and if a Methods section is not available, in a suitable alternative part) of the manuscript."[7]
Elsevier	"Authors should disclose the use of AI Tools for manuscript preparation in a separate AI declaration statement in their manuscript upon submission and a statement will appear in the published work. Authors should document their use of AI, including the name of the AI Tool used, the purpose of the use, and the extent of their oversight."[8]
Emerald	"Copywriting (creating, drafting, or writing) any part of a submission using generative AI tools and technology to generate new material is not permitted."[9]
Wiley	"Authors may only use AI Technology as a companion to their writing process, not a replacement. As always, authors must take full responsibility for the accuracy of all content, and verify that all claims, citations, and analyses align with their expertise and research. Before including AI-generated content in a journal submission, authors must carefully review it to ensure the final work reflects their expertise, voice, and originality while adhering to Wiley's ethical and editorial standards."[10]

degree or credit shall be put in abeyance for a period recommended by the Institutional Academic Integrity Panel and approved by Head of the Institution. Rule 12.2 mentions the penalties in case of plagiarised publications, and the rule seems to apply to faculties and impacts their increments and whether they get to supervise students pursuing research degrees. A note to Rule 12.2 mentions the possibility of suspension or termination in case Level 3 offence is repeated where similarity is above 60%.

Thus, plagiarism has legal ramifications in India. There can be other legal ramifications if copyright infringement case is made in addition to plagiarism case. Thus, it is important to discuss briefly India's stance on copyright vis-à-vis plagiarism especially in the case of AI.

Copyright Infringement vis-a-vis Plagiarism in India and the Role of AI

The Indian Copyright Act, 1957 states in Section 13 that copyright subsists in original literary work. For a work to be original, it must not be copied from another work and must originate from the author of the work, and the work must be a result of the authors' skill.[13] The Indian law requirement for originality is also exercise of skill and judgment and minimum degree of creativity.[14] The Copyright Act defines 'author' in Section 2(d)(i) as "in relation to a literary or dramatic work, the author of the work" and in section 2(d)(vi) as "in relation to any literary, dramatic, musical or artistic work which is computer-generated, the person who causes the work to be created."[15] Section 17 of the Copyright Act states that "the author of a work shall be the first owner of the copyright therein."[15]Section 51 of the Copyright Act states that copyright shall be infringed on unlicensed doing of anything that is the exclusive right of the copyright owner.

As per Section 14 of the Copyright Act, the exclusive rights include "to reproduce the work in any material form".[15] Thus, copyright would be infringed by a person when the person without a license reproduces a work of another person in which copyright subsists. While plagiarism is concerned with passing off someone else's' work as ones' own, copyright infringement is concerned with unlicensed use of work in which copyright subsists.

There could be situations where plagiarism and copyright infringement occur simultaneously. For instance, when people copy other works and do not acknowledge the source, both copyright infringement and plagiarism have occurred. However, there is a 'de minimis' requirement which is still evolving in intellectual property matters in India16 where if the infringement is insignificant or inconsequential, it is a valid legal defence. Therefore, if a student verbatim copies a paragraph from a copyrighted work without acknowledgement, it will constitute plagiarism but only 'de minimis' copyright infringement. Further, if a student copies verbatim a large part of a copyrighted work with acknowledgement but without permission, it could constitute copyright infringement but not plagiarism.

The role of AI such as ChatGPT needs to be ascertained from the perspective of copyright and plagiarism. As per section 2(d)(vi) of the Indian Copyright Act, author of a computer-generated work is the person who causes the work to be created. This provision indicates that the law sees the author of work to be a person and not the AI itself. However, the issue of who should be the author and owner of copyright in AI generated work is a complicated issue.[17] The present-day practicality is that AI cannot be owner of copyright because AI cannot exercise copyright by itself like a human would.

An interesting question nonetheless is whether AI can be recognised as an author in a work while the owner of copyright can be a non-AI entity such as humans. Here, the position of the publisher Taylor & Francis is that AI should not be recognised as the author of a research work because AI cannot take responsibility for the content.[18] A human should be recognised as the author because they can be held accountable for the originality, validity and integrity of the content.[18] A key argument for why human should be considered as the only author is to hold the human accountable for their manuscript and for the human not to disclaim liability by passing the blame onto AI.

The question of whether AI is essentially a co-author, tool, or content creator is an important question. There is an important reason why AI must be seen merely as a tool even if it works beyond what a tool is understood to do. The reason is that AI such as ChatGPT is constantly evolving, and therefore, it would answer the same prompt differently at different times. Therefore, it would be difficult to attribute something to the AI itself which does not seem to have a non-transient attributable opinion.

Whether a human should be copyright author of work created using AI is also another complicated question. Essentially, there must be some creative involvement[19] of the human in the process of creation where they fulfil the expected standards of copyright

authorship in the work. However, answering thisparticular question is also not the focus of this paper.

The focus of this paper is to answer the research question- how should plagiarism policies develop in the age of ChatGPT? This paper believes that there is a need to develop academic integrity both in spirit and practice.[20] Further, there is a need to recognize AI use in research as a serious issue, as reflected in a study of 200 doctoral dissertations where 6% of dissertations from 2023 were positive for AI generated text.[21]

The paper takes a thoughtful approach which is to describe students' perspective in navigating plagiarism softwares and the papersuggests the way forward by proposing an AI plagiarism taxonomy and AI plagiarism policy. The paper relies on the methodology of doctrinal research and has aspects of reflections of ethnographical[22] and autoethnographical[23] nature. The research is carried out by referring to literature on plagiarism generally, and other literature. The information on students' progression in adjusting with plagiarism softwares from the years pre-ChatGPT to years post introduction of ChatGPT is gathered both from existing literature and reflections of the author and his interactions with students at University in India. The conceptual framework of the thesis essentially consists of the concept of plagiarism. The concept is considered while analysing the evolution of student learning from pre-pandemic times to post-pandemic times, and from pre-ChatGPT times to present day.

The justification for relying on the above methodology is that there is a need to longitudinally consider the students' perspective in adjusting to the plagiarism norms to better understand how plagiarism policies need to be designed. There is a need to consider the shifting perspective in a longitudinal manner so that the evolution is traced from pre-ChatGPT days to days post introduction of ChatGPT. This reflective aspect of research is justified because the author has been part of university system since 2014-present which includes before and after the introduction of ChatGPT and has therefore seen the transition in student learning and handling plagiarism softwares in the appropriate longitudinal timeframe. The findings are also validatedat some places from existing literature.

Insights into Student Progression in Dealing with Plagiarism Software Days Pre-ChatGPT to Days Post Introduction of ChatGPT

The paper now describes insights into learning in the pre-ChatGPT age, and how it was possible to cheat the similarity software in that age. This is followed by a discussion onhow things changed with introduction of ChatGPT, and how cheating has taken the form of evading AI use detection by software. Considering these insights, the paper envisions the future idea of plagiarism for protecting honesty and progress in education. These insights are presented in a narrative form from the authors' present recollection of the past eleven years at university.

Learning by Avoiding Similarity in the Pre-ChatGPT Age

Years back, in an age before ChatGPT, the major concern of students was to avoid similarity levels exceeding the prescribed norms. Thus, the studentstook effort in ensuring not to merely copy text and paste it as their own. Similarity check was a deterrent from passing off someone else's work as one's own work. As the software was not easily accessible to students, students had to work to write by themselves.

In the beginning, students would read materials and attempt to paraphrase them in their own words. This would help students deal with the fear of their work being highlighted for similarity. Thus, around that time, similarity software was not just a great deterrent from verbatim copying but a facilitator of original writing. To avoid writing in verbatim language, students were pushed to paraphrase and eventually learn to write in their own words. While similarity checks were not a definitive metric of plagiarism,[24] verbatim copying was avoided, and students wrote by themselves.

Writing by oneself is not just an exercise in avoiding similarity flagging, but an exercise in thinking about what one wants to say and expressing it in one's own unique way. In the process of learning to avoid similarity, one learns to find one's voice by expressing one's thoughts in one's language.

Here, it is important to explain the process of learning. For a student who is not used to writing by themselves, the first step is learning to paraphrase in a way that escapes similarity. The process is not as simple as moving a few words around and changing the tense of the sentence. While paraphrasing, one must learn how to make new sentences to say something with the same or similar meaning as the sentence they are paraphrasing from. However, when words and their sequence change, so does the expressed meaning. In deciding what would be an appropriate paraphrase, one needs to think about whether the sentence conveys a meaning that they agree with. Thus, in creating paraphrased language, one engages in thought.

When a student has gone through the process of paraphrasing several times, they start to learn to write by themselves. It would involve thinking about ideas and being able to state them without having to look at another source to paraphrase from. Initially, the process would involve stating things that they have learnt about. With time and experience, students would be able to generate new knowledge and ideas while expressing them through writing.

The student would create new meaning through the process of writing. Language would then flow from them seamlessly without worry about plagiarism because of similarity. This level of expertise could arise for many students only when they have gone through the learning curve described above.

Cheating the Similarity Software in the Pre-ChatGPT Age

The hurdle of similarity software put honest students through the process of learning to write by themselves and to think by themselves. However, there were dishonest means of bypassing this learning curve by bypassing Turnitin.[25] Students started to buy access to the software used by universities, and to remove similarity in their documents by different means. One such means is to place invisible quotations in the document at places where there is copying of verbatim text. Another is to place invisible fillers in between words that would change the sentence to something not being a verbatim copy.[25] Thus, plagiarised documents would pass off as not being similar. Unfortunately, the focus of academic institutions was mostly on ensuring similarity levels below the prescribed norms[26] and most students had learnt how to evade the software and the human check.

Another way of escaping paraphrasing was to change the poorly paraphrased parts of the document to better paraphrases that would escape the similarity software. In fact, there has been software such as QuillBot[27] which would paraphrase the text to help escape similarity.[28] However, the software would often use synonyms to paraphrase, changing the intended meaning of the text. While it could escape similarity, the problem was that the text read odd because of the unusual words.[29] Again, students could paraphrase the text derived from Quillbot to make it sound more natural. Thus, students found an easy way of escaping similarity by using paraphrasing software.

The Change on Introduction of ChatGPT

The advent of ChatGPT was a major change because students could not only paraphrase existing text but also create new text by itself on any given prompt. Students could delegate the writing to this tool. ChatGPT did not copy verbatim text from existing sources but created new language by itself. Thus, the text produced by ChatGPT would not be flagged for similarity by software. Even if the output of ChatGPT was possibly verbatim plagiarized[30], it could be paraphrased.

Thus, students did not need to go through the process of writing by themselves. Now, students had to intelligently prompt ChatGPT, requesting information on a pointed question. ChatGPT would respond in text relevant to the question. With further intelligent prompting, the text could be further edited, paraphrased, shortened, or otherwise modified.

While ChatGPT may not answer complicated questions in the peculiar way in which the teacher expects the student to answer, the response of ChatGPT would be logical, nonetheless. In several instances, ChatGPT would be able to respond to questions in acceptable ways. In fact, there may be instances where the response is better or comparable to response of a human.[31] Besides, the knowledge of prompt engineering can help get better responses from ChatGPT.[32] Importantly, it is possible to generate a response using ChatGPT that would be of a quality that would be acceptable.

Thus, ChatGPT challenged the notion of plagiarism as text similarity to existing sources. ChatGPT also made it possible to generate text on any topic, therefore making it possible to create text on any assignment that was given to the student. The students could now quickly and easily generate their projects, assignments or any other task.

The students who began university while using ChatGPT missed the pre-ChatGPT age learning curve of struggling to write by oneself, by first learning to paraphrase to avoid similarity and then struggling to ideate through writing. The ChatGPT era students could use ChatGPT to get text written on any idea that they had. The student had to think of a relevant prompt which would lead to a relevant output. The learning curve of writing by oneself was lost in this transition into the age of ChatGPT.

Cheating in the Age of AI Detection by Plagiarism Software

Thus, the concept of plagiarism in the age of ChatGPT could not, therefore, be limited only to the idea of similarity. Plagiarism software thus developed new features for detecting whether text was generated using AI.[33] Consequently, if plagiarism policies disallow text generated by AI to be used in academic

work[5], there would be a need for students to submit text that was not a verbatim copy of the AI output.

Now, the students are faced with a problem similar to the pre-ChatGPT age. Students now need to paraphrase not just what they are taking from published work but also need to paraphrase ChatGPT output to avoid detection of AI use. Students could again use software such as Quillbot or other software that paraphrases the AI text or published text. There could be software that would paraphrase or rework the text in a manner that the plagiarism software does not highlight for AI use.[34]

The dishonest student again had a way – the text can be generated by ChatGPT by a prompt and can be run through AI detection remover software, which makes it difficult to detect that the text was generated by AI. The text will also not be highlighted for similarity as it is not a verbatim copy of published material. On the other hand, the average honest student would use ChatGPT to generate the relevant text and may themselves paraphrase the text.

If students don't resort to software to paraphrase AI-generated or published work, the struggle is to generate paraphrases by themselves. This again would help them undergo the learning curve of having to make sense of information and write in their own language. This would gradually push them towards expressing ideas in their own language and further to ideate through their own language.

In any case, it would be difficult to imagine students not using ChatGPT[35] and instead writing all by themselves. ChatGPT often writes in better language than humans[36] and has the intelligence to quickly respond to specific questions whose answers would be time consuming to find by oneself.

Envisioning the Future of Education and the Role of Plagiarism Software

In the age of AI, what ultimately should be the learning curve for students? Should students have to learn to write by themselves without verbatim copying from AI? Or should it be alright to use AI to generate text that can be verbatim copied into academic work?

Here, it is important to consider that the use of AI in education is inevitable[37] in the coming times, as AI will get better. Thus, it is important to devise policies considering that students would use AI[38] to look for information and answers to their assignments. Policies cannot assume that students should not use ChatGPT in todays' AI age.

However, what is the extent to which the use of AI must be permitted? It is important to understand the functionality available with plagiarism software. The software allows for the detection of AI use. The detection of AI use is not fully accurate, and there can be false positives and false negatives.[39] It is also important to note that there is other software available to paraphrase and escape AI detection. Therefore, it is difficult to say whether a person has written a text wholly by themselves or has merely paraphrased the text from existing sources. It raises important questions about whether academic work must be allowed to be presented when it has not been written oneself.

Given the limitations of plagiarism software, how should we detect and enforce plagiarism? How should we avoid dishonest students from passing the education system without having understood the content of what they are submitting?Can we ensure that the student has understood the meaning of what they have written? One needs to consider that the education system often trains students to memorise content and replicate it. It leads students to replicate what is required without understanding.The disconnect between words and student understanding is amplified when ChatGPT output is used. This is because the words picked up from ChatGPT are not the words of the student.

It is also important to consider that the age of ChatGPT may have changed the learning curve for students. Now, students would learn to use ChatGPT to derive answers that are relevant. The learning involved is to intelligently prompt to get to a desired answer. If education now means looking for new answers to unanswered questions, the role of academics would be to educate students to find answers through all the available tools.

Should the focus still be on writing the words by oneself, or should the focus be on deriving intelligent answers to difficult questions? There is a need for a reimagination of plagiarism as something more than similarity to existing sources or AI output. It requires reimagining plagiarism to put students through a learning curve that is meaningful.

While plagiarism checks should continue to check for similarity to existing materials and to AI output, plagiarism software also needs to evolve to identify instances where text has been merely paraphrased to avoid similarity. Can the plagiarism software also detect whether ideas are original? It would be difficult for plagiarism software to detect if text contains ideas

from existing text. This is because it would be difficult for plagiarism software to capture all the ideas that exist in the literature. There are many ideas in the literature, and there are different ways in which they have been expressed. The same ideas can be expressed in multiple ways in different sequences of words. To identify if a text has the same ideas, it requires not just checking for paraphrasing but for checking whether the ideas are rephrased. The ideas may be rephrased in a language that is not a simple paraphrase.

Further, while each sentence can be evaluated for whether it involves simple paraphrasing, it would be difficult for software to understand whether an original idea emerges from the overall text. Here, human experts could identify whether an original idea emerges from the entire text.[40]

Policy Proposal - How to Determine Plagiarism?

How should plagiarism then be identified by humans? The first step to identifying plagiarism is to check for similarity and AI use. However, there is also a need to look at the text itself. Here, there is a need to ensure proper citation so that the information that is being presented is from a valid source and not just fake information made up by ChatGPT.[41] It is also important to ensure that the references are relevant, and the students do not reference any source which does not contain the stated information. The references should ideally also be authentic and not unscholarly works. Thus, plagiarism would be determined at two levels – one is to avoid similarity and AI flagging beyond a particular limit, and the second is to ensure that sentences that are paraphrased have relevant references especially where they talk about information that is not common knowledge.

The Different Kinds of AI Plagiarism –A Taxonomy

Drawing on the different kinds of plagiarism recognised by Harvard University,[4] and modifying them for the case of AI, the following is the proposed AI plagiarism taxonomy –

(i) Verbatim copying plagiarism - Use of AI output verbatim without quotation or acknowledgement and without citations to the original sources of information for each sentence inside the quotation[40]

(ii) Uncited paraphrasing plagiarism - Paraphrasing AI output without acknowledgement and citations to original sources of information

(iii) AI idea plagiarism - Use of AI to generate ideas which are presented as ones' own original ideas

(iv) Language change plagiarism - Use of AI for language enhancement without acknowledgement of the use of AI such as Grammarly

(v) Misrepresented citation AI plagiarism - Use of AI to generate text, which is used verbatim or paraphrased, and the citation for the sentences is misrepresenting the original source of information

(vi) Prohibited use of AI plagiarism - Use of AI in a way that it has been prohibited to use by the instructor or institution. For instance, use of AI in situations when some institutions have prohibited use of AI in evaluated assessments.

The following can be considered as exceptions to AI plagiarism -

(i) Common knowledge and reasonable expectation - where the idea is common knowledge and has therefore not been cited. And where it can be reasonably be expected that citation is not required because it is commonly understood idea of the field of publication. However, if AI text is copied verbatim, there must be quotations and acknowledgement. If AI text is paraphrased, there must be acknowledgement of use of AI.

(ii) Lack of Intention or recklessness - where the person has unintentionally or without reckless behaviour missed citing the original source or acknowledging AI use. However, when the need to cite and acknowledge is mentioned to the student, they must appropriately revise the draft.

AI Plagiarism Policy for Academic Institutions – Tier 1 of the Policy

The first level of scrutiny can be conducted by drawing a framework based on the UGC Plagiarism Regulations, 2018,[3] where the following mechanism can be used to identify and regulate AI plagiarism -

(i) AI usage Level 1 - 0-10% No penalty

(ii) AI usage Level 2 - 10 -40% Submit revised draft within stipulated time which should have AI percentage less than 10%

(iii) AI usage Level 3 - 40-60% Change of topic and need to submit project again

(iv) AI usage Level 4 - Above 60% Need to wait for six months until a new submission can be made on a different topic

(v) Repeat Offences at Level 4 - Suspension

If the assignment instructions prohibited AI usage, and the software identifies Level 3 or 4 AI usage, it should result in fail grade. If the assignment instructions prohibited AI usage, and the software identifies Level 1 or 2 AI usage, a re-submission should be required after a gap.

While there may be concerns that even when language editing AI is used, it can lead to highlighting of document for AI usage, there must be an option in the plagiarism software to enable or disable check for language editing using AI. When this option is disabled, AI usage should not be highlighted even if a person has edited language using AI software. If the software is unable to disable AI usage highlighting for language editing using AI, the instructors can prohibit language editing using AI to be used.

The above framework would ensure that students at the least need to paraphrase the content they derive from AI. The operationalization of the framework would depend on the robustness with which software can identify AI generated text.

How to Realistically Integrate Human Oversight in AI Plagiarism Detection – Tier 2 of the Policy

A further investigation needs to be carried out in addition to the above scrutiny. The investigation must be carried out by the examiner who is likely knowledgeable in the subject to indicate to the student -

(i) Sentences which are stating a fact or idea that is not common knowledge and where citation is expected.

(ii) Sentences which are stating a fact or idea which has been cited; however, the citation is misrepresenting the source (here the examiner needs to ensure situations are avoided where a source is cited which states exactly the opposite view of what it has been cited for OR situations are avoided where appropriate source is not cited as per usual citing norms).

(iii) Missing acknowledgement of use of AI in idea generation, text generation, paraphrasing and other means.

The examiner should themselves highlight the sentences in the document with such errors and return the document for corrections. On successful corrections with a maximum of two attempts, the plagiarism issue can be considered addressed. To make this process simpler for examiners, software could be developed which highlight such sentences by themselves. The software would need to check for instances of misrepresenting sources, lack of citation where it must be required, and lack of acknowledgement of AI for the kind of usage that has been employed.

Here, it must be noted that unintentional plagiarism is possible where a person has arrived at an idea without coming across a publication which has already arrived at that idea. In such cases, software should also be developed to identify where an idea in a submitted text is originating from.[41] In the case of research at the level of research degrees and not necessarily in the case of scholarly research publications, the originality lies in the process through which an answer has been derived, and not necessarily in the answer itself. Therefore, idea overlaps, i.e. instances where the idea in a submitted work is similar to idea of other work, must be actionable only where there is intentional and reckless disregard for acknowledgement despite the opportunity to do so.

Practical Concerns – Legal Ramifications and State of Software

Lastly, the legal ramifications of accepting AI content with or without attribution must be considered. First, since the copyright over work generated by AI is an unsettled issue, the resulting work would be challenged to not be copyrightable. Further, if the AI generated work does not cite sources where the information exists, there may be challenges to the work in terms of academic honesty. The research degrees that have been awarded and were created using AI could be in question. If AI was used to generate submissions, the scholar has bypassed the learning curve and effort for which the research degree is awarded.

It is also important to ask if there are feasible softwares to detect paraphrased ideas? The software Turnitin claims that it can help "identify text that might be prepared by a generative AI tool such as large-language models, chatbots, word spinners, and bypasser tools."[42] However, it mentions that, "Our AI writing detection model may not always be accurate (it may misidentify human-written, AI-generated, and AI-generated and AI-paraphrased text), so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's

application of its specific academic policies to determine whether any academic misconduct has occurred."[42]Therefore, in the present circumstances where the AI plagiarism identification tools have not progressed, there is a need to give the examiner the reasoned flexibility in deciding a lower level of penalty even when Level 2-4 AI usage is detected. However, the second level of investigation must continue where sentences are identified which are not cited and where AI use is not acknowledged. Further, the website also mentions that scores below 20% are no longer surfaced. This is because "[f]alse positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models. To avoid potential incidence of false positives, no score or highlights are attributed for AI detection scores in the 1% to 19% range."[42] Therefore, at present, the percentage of AI usage for each Level 1-4 can be suitable adjusted depending on the software capability. For now, Level 1 AI usage can be 0-20%, Level 2 can be 20-40%, Level 3 can be 40-60% and Level 4 can be above 60%.

Conclusion

There is a need for academic institutions to shift away from the understanding of plagiarism as similarity alone.[24] Policies need to consider that the process of learning has changed in the age of large language models such as ChatGPT. Therefore, while similarity and AI use must be allowed only to permissible limits, there is a need forfurther policy to ensure honesty in learning. There is a need to check the entire document for ensuring relevant references for the sentences that are not original but contain information.

The students would need to not just learn to paraphrase to avoid similarity and AI flagging but also learn to find sources that state the information that they are presenting in their work.Students would be pushed towards the learning curve of familiarizing themselves with existing literature and being able to place their text submission in the discourse of the relevant literature. Thus, students should be able to find relevant sources, which would require them to look at different papers. This would also help them see, over time, how their work is similar or dissimilar to existing work. They would be able to place their work in the larger context.

It would set a learning curve that supports honesty and quality in work. It would facilitate the progress of science by making students connect their ideas to existing literature. Research-oriented thinking would be incentivised by making students connect the dots in the existing literature and what they are stating. It would lead to a movement of greater awareness of all the research that has happened, and what is unsaid in the literature. This would help create a movement towards novel and original research work.

References

1 Eck N, Using plagiarism detection software: The other side of the Coin, *Academic Integrity*, 435.

2 Van der Sloot B,Regulating the synthetic society: Generative AI, legal questions, and societal challenges, *Bloomsbury Academic*, (2024) 296.

3 University Grants Commission (Promotion of Academic Integrity and Prevention of Plagiarism in Higher Educational Institutions) Regulations, 2018.

4 Harvard University, Harvard guide to using sources: Avoiding plagiarism,https://usingsources.fas.harvard.edu/avoiding-plagiarism-0(accessed on 22 September 2025).

5 Oxford University, Plagiarism,https://www.ox.ac.uk/students/academic/guidance/skills/plagiarism (accessed on 22 September 2025).

6 SAGE, Editorial Policies, Artificial Intelligence Policy, https://www.sagepub.com/journals/editorial-policies/artificial-intelligence-policy (accessed on 22 September 2025).

7 Springer, Artificial Intelligence,https://www.springer.com/in/editorial-policies/artificial-intelligence--ai-/25428500?srsltid=AfmBOopyFPcWYCPsR2e61i0Lq6lePU0SyR2HjBhJe8645LnpPtzyqHJi (accessed on 22 September 2025).

8 Elsevier, Generative AI policies for journals,https://www.elsevier.com/en-in/about/policies-and-standards/generative-ai-policies-for-journals (accessed on 22 September 2025).

9 Emerald, Publishing Ethics,https://www.emeraldgrouppublishing.com/publish-with-us/ethics-integrity/research-publishing-ethics#ai (accessed on 22 September 2025).

10 Wiley Author Services, Best practice guidelines onresearch integrity and publishing ethics, https://authorservices.wiley.com/ethics-guidelines/index.html#22 (accessed on 22 September 2025).

11 WIPO, Generative AI navigating intellectual property, https://www.wipo.int/documents/d/frontier-technologies/docs-en-pdf-generative-ai-factsheet.pdf(accessed 22 September 2025).

12 TOI Tech Desk. Student goes to court after OP Jindal Global Law School fails him over 'AI-generated' answers. https://timesofindia.indiatimes.com/technology/tech-news/student-goes-to-court-after-op-jindal-global-law-school-fails-him-over-ai-generated-answers/articleshow/114949856.cms(accessed 22 September 2025).

13 *Rupendra Kashyap* v *Jiwan Publishing House* 1996 PTC 439 Del.

14 Sharma, S, Originality for Copyright Protection in Literary Works: After *EBC* v *DB Modak*, *Journal of Intellectual Property Rights*, 27(4) (2022) 266.

15 Indian Copyright Act, 1957.

16 Lodha N&Nandkeolyar M, Defense of 'De Minimis' in IP matters in India. https://legalblogs.wolterskluwer.com/copyright-blog/defense-of-de-minimis-in-ip-matters-in-india/.

17 Naithani P, Issues of authorship and ownership in work created by artificial intelligence-Indian copyright law perspective, *NTUT Journal of Intellectual Property Law and Management*, 11(1) (2022) 1.

18 Taylor & Francis, AI Policy,https://taylorandfrancis.com/our-policies/ai-policy/(accessed on 22 September 2025).

19 Liu W & Huang W, Authorship in Human-AI collaborative creation: A creative control theory perspective, *Computer Law & Security Review*, 57 (2025) 106139.

20 Kaktiņš L, Does Turnitin support the development of international students' academic integrity?, *Ethics and Education*, 14 (4) (2019) 430, doi: 10.1080/17449642.2019.1660946.

21 Murch H, Worley M & Volk F, A comparative analysis of the prevalence of artificial intelligence and plagiarism in doctoral dissertations using Turnitin, *Journal of Academic Ethics*, (2025) 1.

22 Gobo G, Ethnography. In*Qualitative research*, Sage, 2011, 15.

23 Ellis C, Adams TE & Bochner A P, Autoethnography: an overview, *Historical social research/Historische sozialforschung*, (2011) 273.

24 Meo S A& Talha M, Turnitin: Is it a text matching or plagiarism detection tool?, *Saudi journal of anaesthesia*, 13 (Suppl 1) (2019) S48-S51.

25 Audrey C, Can Students Trick Turnitin?, *Turnitin.com blog*.

26 University Grants Commission (Promotion Of Academic Integrity And Prevention Of Plagiarism In Higher Educational Institutions) Regulations, 2018 New Delhi, the 23rd July, 2018 F. 1-18/2010(CPP-II).

27 Paraphrasing tool, QuillBot.

28 Rogerson A M& McCarthy G, Using Internet based paraphrasing tools: Original work, patchwriting or facilitated plagiarism?, *International Journal for Educational Integrity*, 13 (2017) 1.

29 Prentice F M & Kinden C E, Paraphrasing tools, language translation tools and plagiarism: an exploratory study 14(1) (2018) *International Journal for Educational Integrity*, 1-16.

30 Lee, Jooyoung, Thai Le, Jinghui Chen, and Dongwon Lee. Do language models plagiarize?. *In Proceedings of the ACM Web Conference 2023*, (2023) 3637.

31 Mostafapour M, Fortier J H, Pacheco K, Murray H& Garber G, *ChatGPT* v *scholars*: a comparative examination of literature reviews conducting by humans and AI, *JMIR Artificial Intellectual*, 3 (2024) e56537.

32 Gao A, Prompt engineering for large language models, *Available at SSRN 4504303*,(2023).

33 Turnitin, Turnitin's AI detector capabilities, *in.turnitin.com.*

34 AI Detection Remover, ByPass AI etc.

35 Sublime J & Ilaria R, Is ChatGPT massively used by students nowadays? A survey on the use of large language models such as ChatGPT in educational settings,*arXiv preprint arXiv:2412.17486*, (2024).

36 Herbold S, Hautli-Janisz A, Heuer U, Kikteva Z & Trautsch A, AI, write an essay for me: A large-scale comparison of human-written versus ChatGPT-generated essays, *arXiv preprint arXiv:2304.14276*, (2023).

37 Krašna M & Bratina T, The use of AI and student population: The change is inevitable, In *2024 13th Mediterranean Conference on Embedded Computing* (MECO) (pp. 1-5), IEEE (2024, June).

38 Rana N K, Generative AI and academic research: A review of the policies from selected HEIs, *Higher Education for the Future*, 12(1) (2025) 97.

39 Dalalah D & Dalalah O M, The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT, *The International Journal of Management Education*, 21 (2) (2023) 100822.

40 Naithani P, Issues of plagiarism and moral right to attribution when using work created by large language models, *Journal of Intellectual Property Rights*, 30 (3) (2025) 376.

41 Safdar M, Nadeem S, Ayesha G, Haisim Y and Muhammad AK, Does ChatGPT generate fake results? Challenges in retrieving content through ChatGPT, *Digital Library Perspectives*, 40 (4) (2024) 668.

42 iThenticate, AI writing detection in the new, enhanced Similarity Report view. https://guides.ithenticate.com/hc/en-us/articles/27835303398541-AI-writing-detection-in-the-new-enhanced-Similarity-Report-view.